GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

VOLUME XVII

1 9 3 5

Issued bi-monthly by the
CLARK UNIVERSITY PRESS
Worcester, Massachusetts

TABLE OF CONTENTS

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

February, 1935
Volume XVII, No. 1

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

ORGANIZATION OF BEHAVIOR IN THE ALBINO RAT

By

ROBERT LADD THORNDIKE

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Philosophy, Columbia University

---

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

ACKNOWLEDGMENTS

The experiment was supervised, from inception to completion, by Professor C. J. Warden of the Animal Laboratory. I am very grateful to Professor Warden for advice and assistance throughout the work. Thanks are due also to Professor H. E. Garrett for suggestions on the statistical treatment. Finally, I acknowledge my debt to my father, Dr. E. L. Thorndike, for many helpful suggestions as to procedure and interpretation.

ROBERT LADD THORNDIKE

COLUMBIA UNIVERSITY
NEW YORK CITY

CONTENTS

I

INTRODUCTION

For a number of years now studies have been conducted upon the organization of behavior in human beings—notably the work of Spearman, Kelly, Thurstone, Garrett, and others. For an even longer time a legion of investigators have been studying animal behavior; but the study of the organization of and relationships within animal behavior is of comparatively recent origin.

Work along this line started as an attempt to determine the reliability of the various instruments—especially the mazes—being used in animal experimentation. In the last fifteen years there have been many of these studies, of varying degrees of experimental and mathematical sophistication, culminating in the recent careful work of Tolman, Stone, Leeper, Heron, and particularly Tryon. The earlier work has been reviewed by Leeper (15), Spence (24), and Tryon (34). Some of it is outlined in Table 1, together with more recent work.

An examination of this table shows us that the earlier studies reported quite low reliabilities. Even these low correlations were partly spurious, inasmuch as different animals were given different numbers of trials. Most of the more recent studies found distinctly higher reliabilities, many of them comparing favorably with those found in human work. The corrected coefficient of .9876 reported by Tryon (34) for 141 rats on a 17-unit multiple-T maze compares very favorably with

TABLE 1
RELIABILITIES OF SCORES IN ANIMAL EXPERIMENTS
A. MAZES

Experimenter	Date	Animals	Maze	Trials correlated	How computed	Correlation		
						(Raw r's)		
Webb (40)	1917	Groups of 6-11 rats	Carr	Mastery	Learn	Trials	—.37 to .13	
					vs.	Errors	—.25 to .56	
					relearn	Time	—.60 to .45	
Hunter (10)	1922	36 and 25 rats	Circular	Mastery	Make	Trials	.31	
					vs.	Total time	.13	
					break	Total dist.	—.02	
						Net time	.36	
			Single-T	Mastery	Make	Time	—.04	
					vs.	Errors	—.07	
					break			
								Median
Hunter (9)	1922	25 rats	Single-T	Mastery	Different		.12 to .59	.32
		36 rats	Circular		tenths		—.08 to .40	.12
		24 rats	Complex circular		correlated		—.12 to .69	.18
Tolman (30)	1924	82 rats	Rectangular	10	Odd	Errors 1-10	.28	
					vs. even	2-10	.30	
					trials	3-10	.38	
						Weighted	.42	
						Corrected	.51	
					Time	Crude	.59	
						Weighted	.53	
						Corrected	.52	
Tolman & Davis (31)	1924	13 rats	Rectangular	3-10	Odd	Crude errors	.27	
					vs.	Corr. errors	.45	
					even	Crude time	.12	
					trials	Corr. time	.63	
			Carr	3-10	Same	Crude errors	.02	
						Corr. errors	.34	
						Crude time	.19	
						Corr. time	.57	

TABLE 1 (*Continued*)

Experimenter	Date	Animals	Maze	Trials correlated	How computed	Correlation	
						(Raw *r*'s)	
Hunter & Randolph (11)	1924	19 rats	T maze	30	Odd vs. even trials Time	.69	
							Median
Hunter & Randolph (12)	1926	32 goats	Single-T	9	Correlated time on single trials	—.01 to .7	.30
					Groups of 3 trials Errors	.25 .06 .32	
Tolman & Nyswander (32)	1927	25 rats	Multiple-choice	3-10	Odd vs. even trials	.48	
					Half vs. half trials	.03	
		24 rats	Circular	3-14	Odd vs. even trials	.56	
					Half vs. half trials	.22	
		42 rats	Carr	3-16	Odd vs. even trials	.60	
					Half vs. half trials	.37	
		15 to 21 rats	3 way	3-25 and 3-10	Odd vs. even trials	.44 to .60	
					Half vs. half trials	.34 to .60	
		18 to 24 rats	Right-left	2-17 and 3-32	Odd vs. even trials	.69 to .88	
					Half vs. half trials	.34 to .59	
		20 to 36 rats	T maze	3-10	Odd vs. even trials	.56 to .72	
					Half vs. half trials	.41 to .65	
		25 rats	Multiple-T maze	4-19	Odd vs. even trials	.83	
					Half vs. half trials	.82	
						(Corrected *r*'s)	
					Errors:		*Median*
Stone & Nyswander (27)	1927	205 rats of different ages	12-unit multiple-T maze	Various groups of the first 30	Odd vs. even trials	.59 to .97	.93
					Odd vs. even blinds	.71 to .96	.87
					Half vs. half blinds	.66 to .96	.85
					Segments of trials	.46 to .97	.86
					Time:		
					Odd vs. even trials	.76 to .97	.90
					Half vs. half trials	.74 to .87	.88

TABLE 1 (*Continued*)

Experimenter	Date	Animals	Maze	Trials correlated	How computed	Correlation	
Stone (25)	1928	201 rats of different ages	Carr	Various groups of the first 20	*Errors:*		*Median*
					Odd vs. even trials	.58 to .87	.72
					Half vs. half blinds	—.04 to .67	.45
					Segments of trials	.12 to .69	.35
					Time:		
					Odd vs. even trials	.25 to .98	.83
					Segments of trials	—.07 to .99	.59
Shirley (22)	1928	29 rats	8-cul maze	Mastery	Odd vs. even trials, Errors	.62	
					Odd vs. even trials, Time	.84	
					Odd vs. even trials minus 1 & 2, Errors	.80	
					Odd vs. even trials minus 1 & 2, Time	.95	
				3–8	Odd vs. even trials, Errors	.73	
					Odd vs. even trials, Time	.94	
Husband (13)	1929	43 rats	Warden, multiple-U	6–15	Odd vs. even trials, Errors	.88	
Miles (17)	1930	58 rats	Elevated multiple-T	Mastery	Odd vs. even trials, errors	.84	
				4–13		.88	
			Alley multiple-T	Mastery	Odd vs. even trials, errors	.86	
				4–13		.84	
Heron (7)	1930	196 rats in various groups	12-unit multiple-T	Two series of 20 each	Same methods used by Stone & Nyswander, with similar results		
					Test vs. retest, with about 200 days interval	(Raw r's) .38 .38	
Tryon (34)	1930	141 rats	17-unit multiple-T	2-19	Odd vs. even trials, Errors	(Corrected r's) .9876	
			20-unit multiple-T	2-19	Same	.9682	
Leeper (15)	1932	202 rats in various groups	12-unit multiple-T	Various parts of 30 trials	*Errors:*	(Raw r's)	*Median*
					Odd vs. even trials	—.44 to .90	.62
					Segments of 10 & 15	—.16 to .65	.46
					Segments of 3 trials	—.27 to .90	.37
					Time:		
					Odd vs. even trials	—.20 to .86	.58
					Segments of 10 & 15	—.21 to .69	.41

Experimenter	Date	Animals	Maze	Trials correlated	How computed	Correlation
Ruch (19)	1930	20 rats	12-unit multiple-T Food reward	2-25	*Errors:*	
					Odd vs. even trials	.68
					Odd vs. even blinds	.83
					Half vs. half blinds	.59
					Half vs. half trials	.46
					Time:	
					Odd vs. even trials	—.01
					Half vs. half trials	—.18
		19 rats	Same, but escape from water	2-25	*Errors:*	
					Odd vs. even trials	.90
					Odd vs. even blinds	.64
					Half vs. half blinds	.51
					Half vs. half trials	.57
					Time:	
					Odd vs. even trials	.65
					Half vs. half trials	.11
Commins, McNemar, & Stone (3)	1932	256 rats	Multiple-T	40	*Errors:*	(Corrected r's)
					Odd vs. even trials	.97
			Elevated multiple-T		Same	.86
			Elevated multiple-U		Same	.86
Tomilin & Stone (3)	1934	136 rats	Warden, multiple-U	20	*Errors:*	
					Odd vs. even trials	.95
			Warden, multiple-U reversed pattern	40	Same	.88
			Elevated multiple-T	16	Same	.71
			Same maze, reversed pattern	10	Same	.50
Dunlap (5)	1933	119 chicks	Warden, multiple-U	8	Odd vs. even trials: Errors	.85
					Time	.89
					Time+errors	.86

TABLE 1 (*Continued*)
B. Tests Other Than Mazes

Experimenter	Date	Animals	Apparatus	Trials correlated	How computed	Correlation	
						Raw	
Hunter (10)	1922	16 rats	T discrimination box	Mastery	Make vs. break	.23	
Heron (6)	1922	28 rats	Inclined-Plane problem box	Mastery	Odd vs. even trials Time	.17	
					Omit trials 1 & 2	.50	
					Omit trials 1-4	.34	
					Omit trials 1-6	.44	
					Omit trials 1-8	.63	
					Odd vs. even tenths Time	.23	
					Omit tenths 1 & 2	.35	
					Omit tenths 1-4	.41	
					Omit tenths 1-6	.46	
					Omit tenths 1-8	.54	
		20 rats	Inclined-plane problem box	12, in two groups of 6, with 60 days between	Trials 1-6 vs. 7-12	—.10	
					Trials 3-6 vs. 9-12	—.22	
							Median
					Individual trials	—.26 to .37	—.01
Hunter & Randolph (11)	1924	15 rats	Sawdust box	30	Odd vs. even trials Time	.25	
Stone (26)	1928	153 rats	Multiple-discrimination	Various fractions of 40	*Errors:*	(Corrected *r*'s)	
							Median
					Odd vs. even trials	.05 to .99	.65
					Segments of trials	—.39 to .98	.53
			Same—reversed cue	Same	Odd vs. even trials	.57 to .92	.84
					Segments of trials	.22 to .93	.77

TABLE 1 (*Continued*)

Experimenter	Date	Animals	Apparatus	Trials correlated	How computed	Correlation
						(Raw r's) ***Median***
Shirley (21)	1928	45 rats	Revolving wheel	20 days	Odd vs. even days	.98
					Sections of 5 days	.83 to .98 .90
				10 days	Odd vs. even days	.98 & .96
						(Corrected r's)
Williams (41)	1929	25 rats	Visual-discrimination	Mastery	Odd vs. even trials Errors	.96
Commins, McNemar, & Stone (3)	1932	155 rats	Triple-plate problem box	11-25 or 11-30	Odd vs. even trials	.67 to .83
		256 rats	Light-discrimination box	40	Odd vs. even trials	.72
Rundquist (20)	1933	Groups of about 20 rats	Revolving wheel	10 days	5 days vs. 5 days	Raw r in the high .80's
						(Corrected r's)
Dunlap (5)	1935	119 chicks	Rotor	2-7	Correlation between *some combination* of odd vs. even trials	.85
			Tunnel	*2 fastest*		.78
			Vocalization	4		.91
			Periscope	2-7		.64
			S maze	1-6		.74
			Prob. box A	1-4		.76
			Prob. box B	1-5		.70
			Alternate stimulus	2-6		.68
			Directional tendency	1-5		.75
Tomilin & Stone (33)	1934	136 rats	Multiple-light-discrimination	30	Odd vs. even trials Errors	.66
			Same—reversed cue	50	Odd vs. even trials Errors	.95

those of any mental tests now available in the human field.

The higher correlations found in the later work may be due in part to more careful technique. They are probably due to a greater extent to improved instruments. The multiple-unit mazes now in use are both more uniform and more difficult than most of the earlier tests, and both these factors are conducive to higher reliability. Tryon lists the following factors as making for higher reliability: (1) more material in test; (2) test-broken individuals; (3) a carefully controlled situation; (4) objective scoring; (5) considerable spread of ability in the group; (6) presence of irrelevant correlated variables (an undesirable source of high reliability); (7) correlation of comparable measures. These factors explain the higher reliabilities found for the recent work, including the very high reliability found by Tryon.

Another source of difference in the size of the reliability coefficients is the variety of ways of computing them. Examination of Table 1 will reveal almost as many ways of computing reliability as there are experimenters. Even the recent workers are not agreed as to the most satisfactory technique. Consider the following six methods: (1) correlating the score on the odd *vs.* even trials; (2) correlating odd *vs.* even blinds; (3) correlating the first half *vs.* second half of the blinds; (4) correlating one segment of trials with another; (5) correlating test with retest; (6) correlating scores on two different mazes. Stone and his coworkers use methods 1, 2, 3, and 4; Tryon favors meth-

od 1; Leeper, method 3; Heron, method 5; Spence, method 6. Apparently no method completely satisfies the two requisites of correlating completely comparable measures and of correlating measures which are free from irrelevant correlated factors. The results from method 1 are almost certainly too high, and those from methods 5 and 6 are probably too low. We cannot say with assurance in which direction the other methods err. But whatever method we may use, we must be fully aware of its limitations.

Intercorrelations between two or more tests were reported as one of the first approaches to the problem of *the reliability of the individual tests.* Later on, intercorrelations were studied for their own sake. Earlier studies were limited to two or three tests; more recently experiments have been extended to include a wider variety of performances. In Table 2 we outline some of the earlier results. We shall consider the more recent work in some detail.

More extensive studies have been published by Commins, McNemar, and Stone (3), by Tomilin and Stone (33), and by Dunlap (5).

Commins, McNemar, and Stone calculated their intercorrelations from data obtained in the course of two other studies. The first study yielded records for several groups of rats on a multiple-T maze, a triple-platform problem box, and the Stone multiple-discrimination box. The reliabilities of these tests were all fairly high. However, the medians of the raw intercorrelations for the six groups for which results were computed were:

TABLE 2

INTERCORRELATIONS

Experimenter	Date	Animals	Tests used	Test reliabilities	Correlation
					(Raw r's)
Bagg (1)	1920	90 mice	Simple maze vs. multiple choice	?	.11
Heron (6)	1922	22 rats	Circular maze vs. problem box	?	.01 to .09
Hunter & Randolph (11)	1924	23 to 37 rats	(1) T maze (2) Straightaway (3) Sawdust box	?	—.16 to .18
Tolman & Davis (31)	1924	13 rats	Rectangular vs. Carr maze	.22 to .77 .04 to .73	Raw .16 to .66 Cor. .6 to 1.0+
Liggett (16)	1925	48 chicks	T maze straightaway	?	—.04 to .19
Williams (41)	1929	Two groups of 25 rats	Visual-discrimination vs. multiple-T maze	.96 .88	Raw .16 & .08 Cor. .17 & .09
Miles (17)	1930	38 rats	Errors, elevated vs. alley multiple-T maze	.84 .86	Raw .50 Cor. .59
Tryon (35)	1931	141 rats	17-unit vs. 20-unit multiple-T maze	.9876 .9682	Raw .77 Cor. .79
Leeper (15)	1932	66 rats	Multiple-T maze vs. mirror image of same maze	Not given for all cases	—.08 to .71 *Median* .30
Shirley (22)	1928	29 rats	Revolving wheel vs. 8-cul-de-sac maze	About .90 Time .97 Errors .89	Raw time —.40 Raw errors —.12 Cor. time —.43 Cor. errors —.13
		17 rats	Same	Same	Raw time —.46 Raw errors —.36 Cor. time —.49 Cor. errors —.40
Rundquist (20)	1933	24 and 31 rats	Revolving wheel vs. maze	Both quite high	Raw: —.36 to .13 Median —.06

Platform box vs. Light-discrimination	.10
Platform box vs. T maze	.02
Light-discrimination vs. T maze	.01

The three tasks seemed to have nothing in common.

The second study gave results for the multiple-T maze, the light box, and two patterns of elevated mazes. *When all the groups were combined*, a population of 256 animals was obtained. For this group, the corrected correlations between the mazes were .56, .65, and .66. The correlations between the light box and the mazes were all close to zero. There appeared to be a community of function among mazes, but this community did not extend to a discrimination habit.

Further light is thrown on this question by the results of Tomilin and Stone (33). Records were obtained for 136 rats on six tests, namely: (1) multiple-U maze; (2) reversed pattern of same; (3) elevated multiple-T maze; (4) reversed pattern of same; (5) multiple-light-discrimination box; and (6) the same with reversed cue. The corrected correlations between mazes ranged from .48 to .86 with a median at .61. The corrected correlation between the two discrimination habits was .66. The median correlation between a maze and a discrimination test was —.02. There seemed to be factors common to the maze habits and factors common to the discrimination habits, but nothing common to the two.

The most extensive study of the interrelations of animal performances is that of Dunlap (5). His subjects were 119 young chicks. These he tested on a variety of simple tests, including several simple maze-

type tests—a straightaway tunnel, an S-shaped maze, and a multiple-U-shaped maze; measures of activity—rotor, periscope, vocalization, problem-box situations; and tests of directional tendency. The reliabilities of the tests were estimated by correlating odd *vs.* even trials and correcting by the Brown-Spearman formula. These reliabilities are included in our Table 1.

The intercorrelations of the different tests were then determined. These were almost all positive and, when different measures from the same test were eliminated, low. The raw correlations between the ten variables which Dunlap used in his final analysis ranged from —.02 to .47 with a mean of .205. The corrected correlations range from —.03 to .65 with a mean somewhere between .25 and .30. Apparently these tests have something in common, though not very much.

Dunlap applied the methods of tetrad analysis to his intercorrelations in order to determine the best pattern of factors to explain them. He feels that the correlations cannot be explained satisfactorily by a pattern of one general factor plus factors specific to single tests, and fits more complicated patterns to the data. He concludes that the complex patterns are very provisional and may well need to be revised with future work, but that the evidence indicated that more than one factor must be postulated to explain the intercorrelation of these abilities in the chick.

My own problem was similar in general to that of Dunlap. I proposed to test a number of albino rats on as wide a variety of problems as possible, to determine the reliabilities of the various scores obtained,

and to find the correlations between them. This table of correlations was then to be examined by the current methods of factor analysis, in an endeavor to trace the relationships existing among the various scores.

II

APPARATUS AND PROCEDURE

The subjects in the experiment were male albino rats, approximately 60 days of age when obtained from the dealer (Breeding and Laboratory Institute in New York City). The rats were run in groups of 18, obtained upon the follow dates: July 17, August 28, and October 23, 1933, and January 3 and February 13, 1934. The homogeneity of the groups is not known, but it is known that the animals were taken from about twenty different litters.

The rats were kept in the laboratory for 12 days before they were started on the experiment proper. During this time they were put through a standard routine of handling and taming, which consisted of about six minutes a day of handling, petting, etc., by the experimenter. Care was taken to keep this treatment as nearly the same as possible for each animal.

Ninety animals started the experimental routine, but there were always some who became conditioned against one or another of the pieces of apparatus and would not run. These failures were almost all in the earliest tests, and these failing animals were discarded, except in one case which will be mentioned later. Our results are based on the complete records of sixty-four animals.

Each animal received the 12-day taming routine described above, and was then put through the standard 39-day experimental routine. The order and time of the different tests was kept the same for the several

TABLE 3

SCHEDULE OF EXPERIMENTATION

Day	8:00 A. M	9:00-10:00 A. M.	4:00 P. M.
1	Maze C preliminary	Jenkins Problem Box preliminary	Maze A prelim.
2	Same	Same	Same
3	Same	Same	Same
4	Maze C trial 1	Same	Same
5	2	Jenkins Problem Box trials 1-2	Maze A trial 1
6	3	3-4	2
7	4	5-6	3
8	5	7-9	4
9	6	10-12	5
10	7	13-15	6
11	8	16-18	7
12	9	19-22	8
13	Maze D trial 1	23-26	9
14	2	27-30	10
15	3	31-35	11
16	4	36-40	Maze B trial 1
17	5	41-45	2
18	6	46-50	3
19	7	51-55	4
20	Latch Box preliminary feeding	56-60	5
21	Same	61-65	6
22	Same	66-70	7
23	Same	71-75	8
24	Latch Box prelim. trials	76-80	9
25	Latch Box trial 1	Latch Box trial 2	Latch Box trial 3
26	4	5	6 Columb.
27	7	8	9 Obstr.
28	10	11	12 Prelim.
29			Columbia Obstr.—hunger
30	"CR" preliminary		"CR" preliminary
31	Same		Same
32	Same		"CR" sound trials 1-10
33	"CR" sound trials 11-20		21-30
34	31-40		41-50
35	51-60		61-70
36	"CR" light trials 1-10		"CR" light trials 11-20
37	21-30		31-40
38	41-50		51-60
39	61-70		71-80

On days 2-15 rats ran in revolving wheel for periods of four hours each day

groups of rats. Table 3 shows how the tasks were temporally distributed.

In general, the animals were experimented with for three or four hours in the morning, and again at the end of the afternoon, about four hours later. The animals were fed on whole wheat bread and milk. They got a little food in the apparatus, when they completed their task; and received their main feeding for ten minutes in the afternoon, at the conclusion of the day's experimentation. Records were obtained for each animal on ten different tests, as follows:

Revolving-wheel activity cage
Warner-Warden maze (2 patterns—A and B)
Elevated T maze (2 patterns—C and D)
Jenkins circular problem box
Latch problem box
Warner's conditioned-response test (two different stimuli)
Columbia obstruction apparatus (new model)—hunger drive

Let us now consider the apparatus, procedure, and scoring for each of these tests in more detail.

Revolving Wheel Activity Cage

A measure of a certain phase of voluntary activity was obtained from the amount that the animal ran in a revolving wheel. The apparatus was made by G. H. Wahmann Mfg. Co., and followed the design of Spaeth. A revolution counter recorded half-turns made by the wheel in either direction.

Due to the limited number of wheels available, it was necessary to run each rat every third day. Each animal was run for five four-hour sessions. The animals were run from eleven or twelve in the morning to three

or four in the afternoon on days 2-15, between the morning and afternoon sessions on the other tests.

The first session for each rat was considered preliminary training.

The score for each animal on this test was the total number of turns of the wheel registered on the second, third, fourth, and fifth sessions.

Warner-Warden Maze

This maze is described in the *Archives of Psychology,* No. 92. Two patterns were used—A and B. These are shown in Figure 1.

The animals were run in Maze A on the afternoons of days 1-15 of the experimental session. On days 1 and 2 they were fed for two minutes in the entrance box of the maze and for about three minutes in the goal box of the maze. On days 3 and 4 the entrance box and goal box were placed next to each other, and the animals were allowed to run from the entrance box to the goal box, where they got a nibble of food. This was repeated five times each day. The sliding cardboard doors of the entrance and goal box were operated during these runs, in order to accustom the rats to them.

On days 5-15 the rats ran through the complete maze from entrance box to goal box, and were allowed to eat in the goal box for about 30 seconds. One trial was run each day.

Days 1-5, including the first trip through the maze, were considered preliminary training.

On days 1 and 2 records were kept of the amount of time that each rat spent in eating, out of the two min-

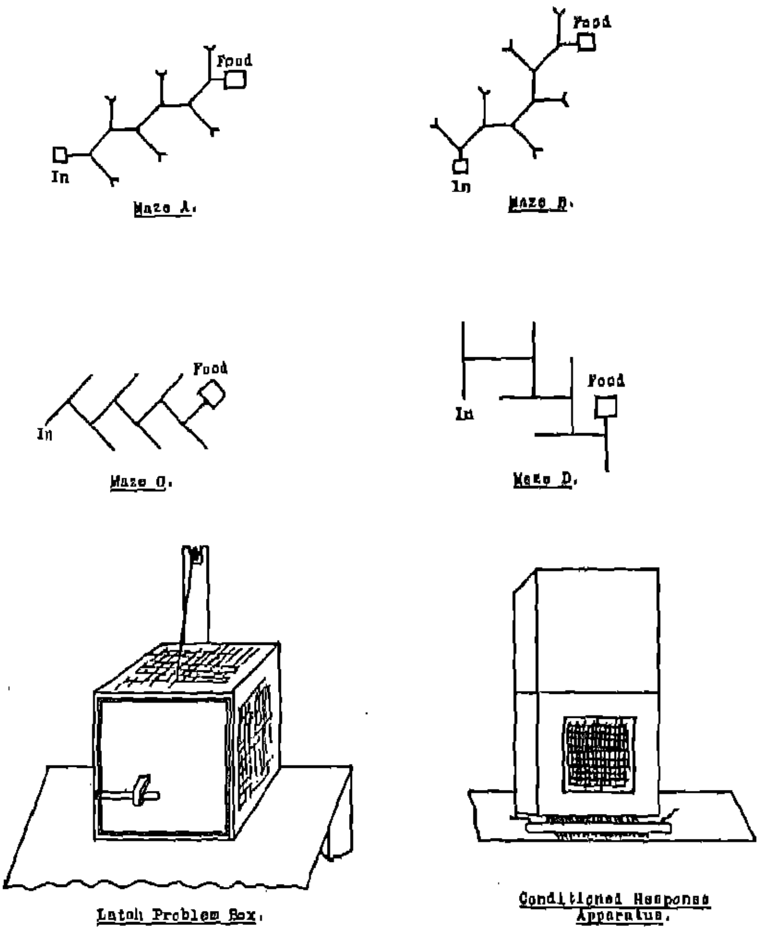


FIGURE 1
DIAGRAMS OF APPARATUS

utes that it was in the entrance box. These records were combined with similar ones from the Jenkins problem box and the Columbia obstruction apparatus

into a score which has been called "Feeding on Preliminaries."

On days 5-15 a complete record was kept of each rat's path through the maze and of his time from entering the maze to entering the goal box. The errors made by the rat in traversing the maze were tabulated in three groups:

"A errors"—entering a blind alley while progressing toward the goal box
"B errors"—entering a blind alley while proceeding in the reverse direction
"C errors"—retracing the true pathway

Each unit of the maze that was incorrectly entered was counted as one error. A tabulation was made of the number of each kind of error for each rat for each trial. The scores on the maze consisted of the sum of the times, or of any single kind of errors, on trials 2-11.

Maze B (second pattern of the Warner-Warden maze) was run in the afternoon on days 16-24. The only preliminary training was the first trial on the maze. Procedure, scoring, etc., were the same as for Maze A. Scores were summed for trials 2-9.

ELEVATED T MAZE

The two patterns of this type of maze were constructed especially for this experiment. In design they resembled somewhat those of Miles (17) and of Dennis (4). They were constructed of planed two-by-fours, fastened together with angle-irons so that the two-inch edge formed the pathway. The rats' pathway was about 15 inches above the level of the maze table.

Two patterns were used—C and D. These are shown in Figure 1.

Maze C was run the first thing in the morning (about 8 o'clock) on days 1-12. On days 1, 2, and 3, the animals were fed in a cage upon the food platform for two minutes. On days 4-12 the rats were placed upon the elevated pathway at the starting end and were allowed to find their way to the food. If they fell off, which happened very rarely and only upon the first trial, they were replaced at the spot from which they fell.

The preliminary training consisted of the three days of feeding and the first run through the maze.

The same records of errors were kept as in the other mazes. In the case of time, however, two separate records were kept. Time was noted from the moment that the rat was put down upon the maze at the starting point to the moment that he left that section of the maze; this has been called "Time to start." The time from leaving the starting section of the maze to reaching the food platform was also recorded; this has been called "Time to run." The scores were the sums of each type of time or errors for trials 2-9.

Maze D was run the first thing on mornings 13-19, with no preliminary training except the first run through the maze. The scores for the maze were the sums on trials 2-7, and were obtained in the same way as for Maze C.

Jenkins Problem Box

The Jenkins problem box used in this experiment is the same as that used by Riess (18) and described by

him. It consists of two concentric circular cages. The inner one is the food compartment. The outer one is the reaction compartment, and contains three plates set into the floor, on some combination of which the animal is required to step. There is an entrance compartment opening into the reaction cage. The door of the entrance compartment is raised to admit the rat to the reaction compartment at the beginning of a trial, and the door of the food compartment is opened when the rat steps on the required plates.

In the present experiment the animals were required to step on any one plate in order to be admitted to the food compartment. Previous workers (23, 18) with this apparatus have used as their first problem the stepping on a certain specific plate. It was not until the experiment was well under way that it was discovered that another procedure had been used before. The other procedure seems better than that which we used, in that it permits the formation of a fairly specific habit. In the problem used here, the rat was rewarded for doing one thing on one trial, and then for doing something quite different on a subsequent trial.

This problem box was run as the second problem on the mornings of days 1-24. The animals were run from 9 or 10 o'clock until 11 or 12.

On days 1 and 2 the animals were fed for 3 minutes in the food compartment. Records were kept of the amount of time spent in eating, as part of our score "Feeding on preliminaries." On days 3 and 4 the door of the food compartment was left open and the animals were allowed to find their way to food, without being

required to step on a plate. Two such trials were given on each of these days. The regular trials of the experiment were given on days 5-24. The number of trials per day started at two and was gradually increased to five.

In all, eighty trials were given. The first thirty trials, together with the training on days 1-4, were considered preliminary training. The scores on this apparatus were based on the records of the last fifty trials.

Four different scores were obtained for this problem box. The first score was a measure of a certain phase of activity in the reaction compartment, namely, the number of quadrants entered. Secondly, a record was kept of the number of perfect trials. A perfect trial was arbitrarily defined as one in which the animal entered no unnecessary quadrants and completed the trial in less than 15 seconds. In the third place, the time between the opening of the entrance compartment door and the entrance of the rat into the reaction compartment was noted. We also kept a record of the time from entrance into the reaction compartment to entrance into the food compartment.

Latch Problem Box

This piece of apparatus was made in the laboratory for this experiment. A diagram of it is given in Figure 1. It consists of a rectangular cage set upon an elevated platform. The wooden door in the front of the cage swings in and up when the catch is released, permitting the animal to get to the food which is placed inside. The latch is a simple wooden bar, set

out from the front of the cage, upon which the animal must push down. When the catch is released, a weight and pulley system swings the door up out of the way.

The rats worked at this problem on the mornings of days 19-24 and on both morning and afternoon of days 25-28. On days 19-23, both before and after the day's run on the Jenkins problem box, each rat was placed on the platform of the latch box and allowed to go into the cage and eat for 15 seconds, the door being left open. On day 24 two trials were given—one in the early morning and one in the late morning—with the latch of the problem box just barely caught, so that any slight jar would release it. This training, together with the first two regular trials, was considered preliminary training.

Twelve trials with the latch fully caught were given on days 25-28, three trials being given each day—one early in the morning, one late in the morning, and one in the afternoon. Each time the rat succeeded in getting into the cage he was permitted to eat for 15 seconds. A time limit of ten minutes was set, and an animal who did not get in in that time was returned to his cage and recorded as having failed for the trial. In this test there were two rats who failed continuously, because they never happened to hit the latch, and not because they refused to work. It was decided not to throw out the complete records of these rats, but to give them a score on this test just worse than the worst of the rats who did not fail.

The only record that was kept for this test was time

to get to the food. The score was the sum of the times on trials 3-12.

CONDITIONED-RESPONSE TESTS

There is room for considerable doubt as to whether the tests to be discussed under this heading should be called conditioned responses at all. If the name is to be applied, it must be with rather a broad definition.

The apparatus used here was the same as that used by Warner (38),[1] as shown in Figure 1. The apparatus was essentially a wooden box with an observation window on one side and a floor made of metal rods. The box was divided in the middle by a low fence of metal bars. The fence was always electrified, and the apparatus was arranged so that either half of the floor could be electrified.

First the animals were trained to jump from one side of the fence to the other when they received an electric shock. Then a buzzer was sounded for ten seconds before the shock was given, in order to train the rats to jump to the buzzer and avoid the shock. Finally a change of illumination—a 100-watt bulb added to a flashlight bulb—was used instead of the buzzer as the signal of the shock.

On the morning of days 30, 31, and 32 and the afternoon of days 30 and 31, ten trials were given with shock alone. The animal was put in the box and presently the shock was administered. The animal was shocked until he managed to get across the fence. He was then

[1] I take this occasion to thank Dr. Warner for permitting me to use his apparatus.

permitted to rest in peace for a minute, after which the shock was again administered until he crossed the fence, and so on for ten trials. These fifty trials and the first ten trials with the buzzer were considered preliminary training.

On the afternoon of days 32, 33, 34, and 35 and the morning of days 33, 34, and 35, ten trials were given with buzzer and shock. On these trials the buzzer was sounded continuously, and if after ten seconds the rat had not crossed the fence, the shock was administered until he did. Then after an interval of from 45 to 70 seconds another trial was given. The animal could and did jump across the fence between trials without incurring any penalty. He could also jump more than once when the buzzer was sounded.

Records were kept of the number of times the animal crossed each time the buzzer was sounded and of the number of times that the animal crossed in the interval between trials. These were used as two separate scores. "Crossings to buzzer" was the number of trials, out of sixty, on which the animal crossed to the buzzer and escaped the shock. "Crossings between buzzers" was the number of crossings that the animal made during the time between trials. These scores were summed for trials 11-70.

On the morning and on the afternoon of days 36-39, ten trials were given with light and shock. The procedure and results were analogous to those for the buzzer. Records were kept for all eighty trials, and gave the two scores "Cross to light" and "Cross between lights."

Columbia Obstruction Apparatus

(Hunger Drive)

The apparatus used in this test was the new model of the Columbia obstruction apparatus, described by Warden (37). This test was given on the afternoons of days 26-29. On days 26 and 27, after the afternoon trial on the latch problem box, each rat was put in the entrance compartment of the obstruction apparatus and was allowed to explore and eat in the apparatus for 3 minutes. No shock was on, and the door in the alley between the entrance and goal compartments was removed. Food was placed in the goal compartment. The amount of time spent in eating was recorded and was combined into the score "Feeding on preliminaries." On day 28 the door was replaced, and the rat was allowed to make 5 crossings from the entrance box to the goal box without shock, getting a bite of food each time. On day 29 the animals, one-day hungry, were allowed four crossings without shock, and on the fifth crossing the shock was turned on. The training up to this point constituted the preliminary training.

After the shock had been turned on, records of the animal's behavior were kept for a ten-minute period. Records were kept for each separate minute, and included approaches (orientations toward the goal compartment), contacts (shocks received), and crossings to the food.

Approximately the same shock was used as was used in the Columbia drive studies (36)—500 volts and .050 milliamperes—but the results are not comparable. In

III

RELIABILITY OF TEST SCORES

Now we come to a consideration of the reliabilities of the various measures taken. In general, these were determined by correlating the sum of the scores on the odd days with the sum of the scores on the even days, and correcting by the Brown-Spearman formula ($R = \frac{2r}{1+r}$). Any deviations from this procedure will be noted in specific cases.

Some recent workers in the field object to this method of computing reliabilities, saying that the correlation is rendered spuriously high by the presence of irrelevant correlated factors in the two halves of the test. This is a very sound criticism. However, the procedure is the most standard one, and any other that we might substitute for it suffers either from the same difficulty or from the difficulty of not correlating comparable things. In the present case we at least know in which direction the error is likely to be, which is more than we can say for some of the other techniques. So we will correlate odd *vs.* even days, remembering that the correlations we get are probably too high.

Let us go through the list of variables, considering the reliability obtained for each.

Revolving Wheel

Variable 1. The total score was the number of turns on days 2-5. The reliability was computed by correlating the number of turns on days 2 and 5 with the number of days 3 and 4. The correlation obtained

was: Raw—.96, Corrected—.98. Apparently four four-hour sessions in a revolving wheel give one a very stable measure of that phase of a rat's activity. This high reliability agrees well with the results of Shirley (21) and Rundquist (20). (See Table 1).

Maze A. (Warner-Warden)

Variables 2, 3, 4, 5. The correlations are for odd *vs.* even trials on trials 2-11.

			Raw	Corrected
Var.	2	"A errors" (Forward into blind alley)	.68	.81
	3	"B errors" (Retracing into blind alley)	.53	.69
	4	"C errors" (Retracing true path)	.78	.88
	5	Total time (Entrance box to goal box)	.73	.84

The reliabilities here are fairly satisfactory. Though they do not measure up to those obtained with more difficult mazes and better controlled conditions by Tryon (34), Stone and Nyswander (27), and Heron (7), they are higher than those reported by earlier maze workers, and higher than the general run of those reported recently by Leeper (15).

Maze B. (Warner-Warden)

Variables 6, 7, 8, 9. The correlations are for odd *vs.* even trials on trials 2-9.

			Raw	Corrected
Var.	6	"A errors"	.63	.77
	7	"B errors"	.65	.79
	8	"C errors"	.69	.81
	9	Total time	.84	.91

It is interesting to note that in these mazes the retracing errors give as high or higher reliabilities than the

forward-going errors. This may afford some confirmation of the experimenter's feeling that "tameness" may enter into the score on a maze of this type and degree of difficulty as much as, or more than, "intelligence."

Maze C. (Elevated Maze)

Variables 10, 11, 12, 13, 14. The correlations are for odd *vs.* even trials on trials 2-9.

		Raw	Corrected
Var. 10	"A errors"	.31	.47
11	"B errors"	.36	.53
12	"C errors"	.32	.49
13	Time to start running	.71	.83
14	Time to run	.31	.47

The reliabilities on this maze are all low, with the exception of "Time to start." The only consistent element of a rat's behavior on the elevated maze was the time that he remained still before beginning to run. Observation of the animals leads the experimenter to believe that this is, in large measure, a manifestation of tameness. The low reliabilities of the other measures may be in part a function of the timidity of the animals on the elevated maze. They seemed more susceptible to disturbance by environmental factors when they were up in an exposed position than when in the alleys of the Warner-Warden maze.

Maze D. (Elevated Maze)

Variables 15, 16 17, 18, 19. The correlations are for odd *vs.* even trials on trials 2-7.

		Raw	Corrected
Var. 15	"A errors"	.50	.67
16	"B errors"	.37	.54

17	"C errors"	.20	.33
18	Time to start	.85	.92
19	Time to run	.38	.55

Again all the reliabilities are low except "Time to start." In the case of this maze, the low reliabilities may be due in part to the fact that the maze was very easy. It is interesting to note that the changed pattern of the elevated maze was learned very readily, while the changed pattern of the alley maze presented a great deal of difficulty. The floor plans were not exactly the same for the alley and elevated mazes and the results may be due entirely to this difference in floor plan. But the possibility is suggested that different sensory cues may be used in the elevated maze. Commins (2) has also observed the readiness with which a second elevated maze pattern is learned.

Jenkins Circular Problem Box

Variables 20, 21, 22, 23. The correlations are for odd *vs.* even days (5 trials per day) for trials 31-80.

		Raw	Corrected
Var. 20	Quadrants entered	.57	.73
21	Perfect trials	.26	.41
22	Time to start from entrance box	.89	.94
23	Time to run	.82	.90

The score for perfect trials, as arbitrarily defined by the experimenter, seems to have little value. Some animals developed a very successful response to the problem which never satisfied the arbitrary definition of a perfect trial, while many perfect trials were apparently achieved by chance. Quadrants entered gives

a fairly consistent measure of a certain fraction of the animal's activity in the problem box, though many phases of his activity do not enter into this score. The time scores give the highest reliabilities, and again "Time to start" seems to be a very reliable measure. Under the circumstances of this experiment, "Time to run" seems to be the best measure of learning the problem.

The high reliability of these measures of the time that elapses before starting in on some performance, after having been put in the apparatus, seems very interesting. We shall see later that the three measures of time to start on different performances correlate quite highly. There seems to be an interesting and not widely advertised phase of the individual animal's make-up which enters in here.

Latch Problem Box

Variable 24. In this problem the animals were run three trials a day, the trials being separated by an interval of from two to four hours. The reliability of the score was obtained by correlating odd *vs.* even trials for trials 3-12. For the time taken to get to the food, the reliability was:

Raw—.77 Corrected—.87

Conditioned Response

Variables 25, 26, 27, 28. In these tests the animals were run ten trials at a time twice a day—in the morning and again in the afternoon. The reliabilities were computed by correlating odd *vs.* even sets of ten trials.

This amounted to correlating the score on the morning trials with the score on the afternoon trials. Reliabilities were obtained as follows:

		Raw	Corrected
Var. 25	Cross to buzzer	.76	.87
26	Cross between buzzers	.82	.90
27	Cross to light	.82	.90
28	Cross between lights	.80	.89

Both the frequency of crossing to the stimulus and the readiness of the crossing response seem to be quite stable features of the animal's behavior. A consideration of the intercorrelations, which we shall give later, suggests that frequency of crossing to the stimulus may be in large measure a function of the readiness of the crossing response.

Columbia Obstruction Apparatus

(Hunger Drive)

Variables 29, 30, 31. This test was run for one continuous period of ten minutes. The reliability was computed by correlating odd *vs.* even minutes. This reliability is the most questionable of any that we have obtained, because it seems hardly conceivable that many of the chance errors affecting the score on one minute would not affect the score on the adjacent minutes. However, the procedure is strictly analogous with that used in correlating odd *vs.* even items in a human test, and we shall use it here, remembering that the obtained reliabilities are almost certainly too high. The obtained reliabilities follow:

		Raw	Corrected
Var. 29	Approaches	.58	.73
30	Contacts	.72	.84
31	Crossings	?	?

In the case of crossings (Var. 31), less than a quarter of the rats crossed the grid even once, because the shock used was too severe. This being the case, a reliability coefficient seemed meaningless, and so was not computed.

The reliabilities here are not very high, and are probably partly spurious, so probably we should not attach much significance to the results on this test. It must be remembered that the shock and procedure which we used were not standard for this apparatus.

Feeding on Preliminaries

Variable 32. The reliability was computed by correlating the amount of time spent in eating on odd *vs.* even days, as there were scores for two days on each piece of apparatus. The reliability was:

Raw .65 Corrected .79

IV

INTERCORRELATION OF TEST SCORES

We have seen to what extent the tests are reliable measures. Now we are interested in finding out how these measures are related to one another. For this purpose we need the correlations between the different variables. The intercorrelations were computed by the Columbia Statistical Bureau. The Bureau makes a practice of computing all correlations to four decimal places, but, inasmuch as the correlations were based on only sixty-four cases and the sigmas of the correlations were of the order of 0.1, the last two places have been discarded as meaningless. The table of intercorrelations for the thirty-two variables is given, to two decimal places, in Table 4. The raw correlations are given on the upper right-hand side, the reliabilities along the diagonal, and the corrected correlations on the lower left-hand side.

In giving these correlations, it has been decided in an arbitrary and common-sense manner what shall be a "good" score for each variable. For example, in all the maze-error scores a good score is a low score, while for the revolving wheel a good score is a high score. Consequently, a positive correlation between the revolving-wheel and a maze-error score indicates that the animal who runs many turns in the revolving wheel tends to make few errors in the maze. The following have been chosen as "good" scores:

Revolving wheel—	large number of turns;
Mazes—	few errors or short time;

TABLE 4
Intercorrelations of 32 Variables*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	98	27	25	25	36	13	07	12	19	26	25	25	25	25	03	06	—01	24
2	30	81	79	88	82	39	45	42	29	53	32	57	46	57	52	39	36	28
3	31	100+	69	87	81	36	46	42	33	56	40	61	46	56	16	37	35	26
4	24	100+	100+	88	84	26	47	38	28	56	54	61	48	58	14	32	29	24
5	40	100	100+	97	84	29	39	36	41	48	39	61	59	69	09	26	20	40
6	15	49	50	52	35	77	74	85	74	20	04	22	11	10	54	35	32	31
7	08	54	63	56	47	94	79	91	79	18	—04	22	18	09	26	37	33	33
8	13	52	56	45	44	100+	100+	87	82	13	—03	23	13	07	19	32	29	23
9	20	54	41	31	47	100+	93	95	97	12	—01	22	27	17	15	16	21	60
10	39	86	99	87	77	32	30	20	19	47	50	72	45	58	40	26	30	18
11	35	49	66	50	59	07	—07	—05	—01	100	53	68	53	68	40	26	05	21
12	33	91	100+	93	100+	35	36	36	33	100+	100+	49	58	81	23	18	20	21
13	28	55	61	56	71	14	22	15	31	72	80	91	87	74	12	08	05	54
14	37	92	98	91	100+	17	15	11	27	100+	100+	100+	100+	47	17	07	10	40
15	04	44	23	18	12	48	35	26	20	71	30	40	16	30	67	38	55	29
16	08	56	60	47	38	54	57	49	23	52	03	35	11	14	63	54	85	10
17	—02	69	72	53	57	62	64	55	38	75	11	50	10	24	100+	100+	33	22
18	25	35	32	27	46	36	39	27	65	27	29	31	62	60	37	15	40	02
19	29	65	67	56	51	61	52	46	35	71	79	69	49	73	90	92	100+	48
20	—29	17	05	04	00	35	17	19	08	—04	—10	—29	—17	—29	29	27	35	04
21	—35	—04	—15	—16	—22	45	27	37	17	—08	—04	—36	—50	—59	37	—16	07	—20
22	36	49	40	34	59	46	33	40	59	32	37	52	47	59	33	33	56	65
23	22	52	37	32	51	58	40	48	54	24	12	23	25	28	37	36	59	52
24	08	16	33	16	18	52	32	25	30	13	03	10	15	14	29	64	39	44
25	57	24	16	12	06	30	11	18	—05	37	32	16	11	04	14	19	12	00
26	35	32	28	17	23	18	00	11	—07	19	15	14	—01	10	—01	18	17	—14
27	21	02	01	—04	—11	17	—04	05	—09	29	31	07	—11	10	09	26	27	—07
28	40	20	29	16	18	39	27	26	16	23	17	22	—07	08	05	19	15	11
29	—38	08	—02	03	12	36	33	35	32	—13	—20	04	25	06	02	10	09	11
30	—02	20	58	17	20	14	18	17	—02	14	00	21	12	19	23	15	32	—16
31	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
32	45	49	37	44	41	24	19	09	22	53	44	56	41	63	43	21	25	46

*Decimals have been omitted.

TABLE 4 (*Continued*)

	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	21	—24	—22	34	21	08	34	33	19	57	—07	—02	—15	38
2	43	13	—02	43	44	14	20	27	01	17	06	17	11	40
3	41	04	—07	32	29	25	13	22	01	23	—01	29	16	28
4	39	05	—10	31	29	14	10	16	—04	14	02	15	11	36
5	55	00	—13	52	45	15	05	20	—09	15	10	17	10	34
6	40	27	24	40	48	26	25	15	14	33	27	12	19	19
7	34	13	15	29	33	26	09	00	—03	19	25	14	20	15
8	31	15	22	35	41	21	15	09	03	22	27	14	18	07
9	25	06	11	55	49	26	—04	—06	—08	15	26	—02	15	18
10	36	—02	—04	22	16	08	24	12	19	15	—08	09	16	32
11	43	—07	—02	26	08	02	22	11	22	12	—13	00	07	29
12	36	—17	—16	35	16	06	10	09	05	14	03	13	10	55
13	33	—14	—30	41	21	13	10	00	—10	—06	20	10	11	33
14	37	—17	—26	39	18	09	02	07	—07	05	03	12	09	38
15	54	20	20	26	28	22	11	—01	07	04	01	17	18	31
16	50	17	—08	23	25	44	13	13	18	13	06	10	14	14
17	52	17	03	32	32	48	07	09	15	08	04	17	18	13
18	54	05	—12	63	47	39	00	—13	—06	10	09	—14	09	39
19	55	17	01	42	38	33	12	11	—02	10	11	08	16	42
20	26	73	51	02	39	07	22	13	25	19	06	05	13	—13
21	02	94	41	—05	24	—07	29	13	26	17	—02	04	11	—21
22	58	02	—08	94	83	21	05	15	00	19	16	08	14	46
23	53	48	39	90	90	27	17	27	10	29	23	16	22	31
24	47	09	—11	23	30	87	05	—04	10	11	03	14	22	20
25	18	28	48	05	19	05	57	54	77	71	—05	11	01	12
26	15	15	21	17	30	—05	61	90	45	64	27	26	04	07
27	—05	31	43	00	11	12	88	50	90	64	—08	00	—06	—02
28	14	24	28	21	33	13	81	72	72	89	14	16	—01	19
29	17	08	—03	19	29	04	—06	34	—10	17	73	49	45	13
30	12	06	07	09	19	17	12	29	00	18	63	84	53	07
31	—	—	—	—	—	—	—	—	—	—	—	—	—	21
32	63	—18	—37	53	37	24	14	08	—02	23	17	08	—	79

Problem boxes— short time, few quadrants entered, and many perfect trials;
CR— many crossings of the fence;
Obstruction apparatus—many approaches, contacts, or crossings.

Let us now examine the table of correlations. In the first place, most of the different measures from the same test (different types of errors, time, etc.) are closely correlated, especially for the mazes. The raw correlations are quite high, and when these are corrected for the unreliability of the tests, many of them become greater than unity. For the Warner-Warden mazes the median raw correlation between different types of scores is about .82, and the median corrected correlation over 1.00. For the elevated mazes the same figures are .54 and .96. Apparently these scores are all measuring much the same thing. Later we shall consider a condensed array of correlations, where several of these scores have been combined into single scores.

When we consider the correlations between different tests, we observe that the great majority of them are positive and that most of them are quite small. Less than 15 per cent of the correlations are negative, and no one of these is significantly different from zero. The mean of the raw correlations is .18 and that of the corrected correlations is .27. This result agrees quite closely with a mean raw correlation of .205 found by Dunlap (5) from his tests on chicks.

The table of correlations shows us quite distinctly that the mazes are more closely related to one another than they are to the other tests, or than the other tests,

in general, are to each other. The average raw correlation between different mazes is about .30, and the average raw correlation between two tests which are not both mazes is about .17. When we use corrected correlations the difference shows up in a more striking way. In this case the two figures are .46 and .19. Mazes have something in common which spreads to a much lesser extent to the other tests used. This was also found to be the case in the work of Commins, McNemar and Stone (3), and of Tomilin and Stone (33). The average intercorrelation of our mazes is somewhat less than they report.

The relationship of activity (Var. 1) and "Feeding on preliminaries" (Var. 32) to the rest of the tests is rather interesting. There is some indication that the rat who is more active and who eats more readily in a new situation learns the maze better, has lower time scores on the problem box, and crosses the fence more readily in the CR apparatus. The same relationship between activity and maze score was reported by Shirley (22), while Rundquist (20) found a negligible relationship.

Variables 13, 18, and 22 are of considerable interest. Each of these measures the time that the animal took to start going in some apparatus. That this trait is fairly consistent from one test to another is indicated by the correlations of .54, .41, and .63 (corrected—.60, .47, .68) between these variables. These measures of time to start correlate with the other maze and problem-box time and error scores, but not as highly as the time and error scores correlate among themselves.

A table of intercorrelations among 32 variables is rather hard to analyze. There is so much detail that it tends to hide any general trends. It therefore seemed desirable to reduce the number of variables. We have seen that the three types of error scores for each maze were measuring much the same thing, so they might well be combined into a single score. We therefore combined the "A," "B," and "C" errors for each maze into a single variable.

In combining the variables it was necessary to adopt some system of weighting. The simplest technique would have been to weight each variable equally, but this would have given undue weight to a variable in which the scores scattered widely. The next simplest thing would have been to weight each variable in inverse proportion to its standard deviation. This would probably have proved perfectly satisfactory, and would have given results that differed only very slightly from those which we obtained by a somewhat different technique.

We shall illustrate the technique which we adopted by an example. Consider variables 2, 3, and 4. The intercorrelations between these variables are $r_{23} = .79$, $r_{24} = .88$, $r_{34} = .87$. We form the table:

	.88	.79	.88	
	.79	*.87*	.87	
	.88	.87	*.88*	
Σr	2.55	2.53	2.63	$\Sigma\Sigma r = 7.71$
				$\dfrac{1}{\Sigma\Sigma r} = .1298$
k_1	.92	.91	.95	$\dfrac{1}{\sqrt{\Sigma\Sigma r}} = .36$

and fit a first factor to the array of correlations by Thurstone's center of gravity method of factor analysis (29). We obtain the factor loadings .92, .91, and .95 for the three variables. The scores on the variables are first reduced to standard scores (divided by their standard deviations), and then given weights equal to these first factor loadings.

The intercorrelations between this new composite variable and the other variables are found from the correlations of the separate component variables by the formula for the correlation of sums. [See Kelly (14, p. 197).] In this particular case, the formula takes the form

$$r_{(Ax+By+Cz,\ w)} = \frac{A.r_{xw}+B.r_{yw}+C.r_{zw}}{\sqrt{A^2+B^2+C^2+2AB.r_{xy}+2AC.r_{xz}+2BC.r_{yz}}}$$

where A, B, and C are the weights (i.e., the factor loadings) of the variables x, y, and z.

In this way, we combined the three error scores for each maze pattern into a single score. We also discarded several variables which were of low or unknown reliability, namely, variables 14, 19, 21, and 31. The correlations among the remaining 20 variables are given in Table 5.

Again we see that most of the correlations are positive, and that the negative ones are quite small. The correlations between mazes are all positive, with a median raw correlation of .36. The correlation between time and error scores for the same maze are quite high—.87 and .83 for the Warner-Warden mazes, and lower for the less reliable elevated mazes. The cor-

TABLE 5

INTERCORRELATIONS: 20 VARIABLES

Variables	1	2 etc.	5	6 etc.	9	10 etc.	13	15 etc.	18	20	22	23	24	25	26	27	28	29	30	32
1		.26	.36	.11	.19	.28	.25	.03	.24	—.24	.34	.21	.08	.34	.33	.19	.37	—.07	—.02	.38
2 etc.			.87	.45	.32	.61	.49	.37	.27	.07	.37	.36	.18	.15	.23	—.01	.19	.02	.21	.56
5				.36	.41	.57	.59	.22	.40	.00	.52	.45	.15	.05	.20	—.09	.15	.09	.17	.34
6 etc.					.83	.16	.15	.38	.30	.19	.36	.43	.26	.17	.08	.05	.26	.28	.14	.14
9						.13	.27	.21	.60	.06	.55	.48	.26	—.04	—.06	—.08	.15	.26	—.02	.18
10 etc.							.60	.27	.23	—.10	.32	.15	.06	.21	.12	.17	.15	—.06	.09	.37
13								.10	.54	—.13	.41	.21	.13	.10	.08	—.10	—.06	.20	.10	.33
15 etc.									.24	.21	.32	.33	.45	.12	.09	.16	.10	.05	.17	.22
18										.03	.63	.47	.39	.00	—.13	—.06	.10	.09	—.14	.39
20											.02	.39	.07	.22	.12	.25	.19	.06	.05	—.13
22												.82	.21	.05	.15	.00	.19	.16	.08	.45
23													.27	.17	.27	.10	.29	.24	.16	.31
24														.05	—.04	.10	.11	.03	.14	.19
25															.54	.77	.71	—.05	.11	.12
26																.45	.64	.27	.26	.07
27																	.64	—.08	.00	—.02
28																		.14	.16	.19
29																			.49	.13
30																				.07
32																				

relations between the different CR variables are quite high—median .64. Most of the other correlations are quite low.

We have further combined the variables, in a more or less arbitrary manner—reducing the number to 13. Variables 2 etc. and 6 etc. have been combined into a single alley-maze error score, variables 5 and 9 into a single alley-maze time score, variables 10 etc. and 15 etc. into a single elevated-maze error score, and variables 13 and 18 into a single elevated-maze time score. Variables 25 and 27 have been combined into a single jump-the-fence-to-stimulus score, and 26 and 28 into a single jump-the-fence-between-stimuli score. Variables 29 and 30 form a single hunger-drive score. The same method of weighting was used as before, but as only pairs of variables were combined, this amounted to giving equal weights to the correlations of each variable. The correlations among the remaining 13 variables are given in Table 6.

TABLE 6

INTERCORRELATIONS: 13 VARIABLES

Var.	1	2,6	5,9	10,15	13,18	20	22	23	24	25,27	26,28	29,30	32
1		.22	.33	.20	.28	—.24	.34	.21	.08	.28	.38	—.05	.38
2,6			.77	.56	.41	.15	.43	.46	.26	.11	.25	.22	.30
5,9				.43	.63	.04	.64	.55	.25	—.05	.14	.17	.31
10,15					.42	.07	.40	.31	.32	.22	.16	.09	.37
13,18						—.06	.59	.39	.30	—.02	—.03	.08	.41
20							.02	.39	.07	.25	.17	.06	—.13
22								.82	.21	.02	.19	.14	.45
23									.27	.14	.31	.23	.31
24										.08	.04	.10	.19
25,27											.68	—.01	.05
26,28												.26	.14
29,30													.12
32													

Examination of this table shows correlations between the various maze scores ranging from .41 to .77, with a median at about .50. The maze scores seem to be related fairly closely to the two Jenkins problem box time scores—the correlations range from .31 to .64 with a median at about .44. Slightly lower correlations appear between the mazes and the latch problem box (.25 to .32), the revolving wheel (.20 to .33), and the preliminary feeding (.30 to .41). The time scores for the Jenkins problem box are related to activity (.34 and .21) and to feeding on preliminaries (.45 and .31). Another interesting relationship is that between activity and CR scores (.28 and .38).

V

ANALYSIS OF INTERCORRELATIONS

The examination of the intercorrelations has so far been qualitative and subjective. Now let us see what we get when we apply current methods of factor analysis to the array of correlations. Of the various possible techniques, we have used Thurstone's center of gravity method (29). Practice has shown that the results from this method approximate quite closely those obtained by Thurstone's (28) and Hotelling's (8) more laborious techniques. We have checked on this for our 13-variable array of correlations by computing the first two factors by Hotelling's method also and comparing them with the results of the simpler method. These results will be compared later. It is sufficient to say that with our sixty-four cases the greatly increased labor of the other techniques did not seem worth while for the 32- and 20- variable arrays.

A factor pattern was first fitted to the original array of correlations between *all 32 variables.* The factor loadings for the first three factors are given in Table 7.

In order to obtain some idea of how many factors are needed to give an adequate fit to the array of correlations, the residuals after removing each factor have been tabulated. If we assume that the true correlation between any two variables is zero, then with our sixty-four cases the standard deviation of this correlation is .125. Working on the assumption that no true correlation remains among the variables after removing one factor, we determine how many of the obtained

TABLE 7

FACTOR LOADINGS OF 32 VARIABLES

	Factor		
Variable	1	2	3
1	.338	.233	—.254
2	.759	.299	.037
3	.730	.308	.061
4	.682	.404	.090
5	.713	.452	.136
6	.656	—.400	.192
7	.611	—.289	.394
8	.608	—.359	.337
9	.563	—.150	.408
10	.579	.350	—.345
11	.432	.486	—.152
12	.606	.517	—.120
13	.515	.532	.101
14	.557	.658	—.039
15	.452	—.096	.043
16	.504	—.226	.222
17	.525	—.252	.271
18	.494	.191	.254
19	.623	.068	.106
20	.199	—.441	—.059
21	.086	—.484	—.201
22	.641	.110	.176
23	.650	—.168	.112
24	.369	—.141	.221
25	.397	—.289	—.769
26	.350	—.216	—.462
27	.251	—.359	—.693
28	.442	—.318	—.563
29	.243	—.270	.233
30	.292	—.232	.100
31	.310	—.215	.193
32	.464	.272	.017
$\frac{\Sigma k^2}{n}$	.272	.112	.088

residuals lie within 1 sigma (.125), 2 sigma (.250), etc., of the assumed true value of zero. This distribution of residuals is then compared with the distribution to be expected by chance. If the distribution of residuals approximates what should be expected by chance, we have some justification in feeling that we have

fitted a sufficient number of factors to the table of correlations.

For the 32-variable table of correlations the results were as follows:

	First residuals %	Second residuals %	Third residuals %	Normal probability %
Within 1 sigma	59.4	73.1	78.4	68.3
2	88.5	93.6	97.8	95.4
3	96.0	98.0	99.6	99.7
4	98.8	99.6	100.0	100.0
5	99.8	100.0		

From this table we can see that the three factors reduce the residual correlation between variables to about what would be expected to arise by chance alone. There may be some doubt as to whether the third factor is needed.

Now let us examine these three factors in order to try to get some understanding of what they might be. In the first place, the first factor has a positive weight in every test, though some of the weights are very low. This shows that every test tends to be correlated positively with the other tests. We must be cautious, however, in interpreting this as a general factor among all the tests; the slight tendency for all the tests to be correlated may have arisen through some slight heterogeneity of conditions among the different groups of animals and nothing more. This first factor is most prominent in the maze scores and in the time scores of the Jenkins problem box. Perhaps it should be considered maze-running ability, rather contaminated with activity, tameness, and possibly hunger.

The second factor seems to be a temporal factor, in that most of the tests that occurred early in the experimental routine have positive weights and those that occurred late in the training have negative weights. Since it discriminates the earlier from the later tests, we might name it a transfer factor.

The third factor seems to discriminate the CR tests from the rest of the tests. They have large negative weights, while most of the rest of the weights are rather small negative or positive ones. The four CR scores were obtained in the same apparatus, at the same part of the schedule, with only slightly different conditions.

TABLE 9

FACTOR LOADINGS OF 20 VARIABLES

Variable	Factor 1	Factor 2	Factor 3
1	.414	.101	.233
2, 3, 4	.684	—.136	.490
5	.690	—.293	.439
6, 7, 8	.612	—.227	—.282
9	.570	—.468	—.342
10, 11, 12	.508	.005	.598
13	.491	—.311	.390
15, 16, 17	.459	—.031	.024
18	.539	—.405	—.033
20	.177	.152	—.361
22	.698	—.332	—.056
23	.716	—.167	—.319
24	.366	—.113	—.080
25	.468	.756	—.040
26	.435	.508	—.090
27	.331	.801	—.135
28	.556	.595	—.220
29	.284	—.243	—.379
30	.278	—.043	—.112
32	.469	—.126	.254
$\frac{\Sigma k^2}{n}$	.259	.136	.086

The apparatus was very different from any of the other pieces used. The motivation was escape-from-shock instead of hunger. Under these conditions it is not surprising that the four CR scores have factors in common which do not extend to the rest of the tests.

Now let us see what we get when we fit the same type of factor pattern to the 20-variable array of correlations. The factor loadings for three factors appear in Table 8.

As in the case of the 32-variable array, we compare the residuals with the correlations to be expected by chance:

	First residuals %	Second residuals %	Third residuals %	Normal probability %
Within 1 sigma	53.7	73.5	81.1	68.3
2	84.7	94.1	97.4	95.4
3	96.3	99.0	100.0	99.7
4	99.5	100.0		100.0

Here again, three factors seem to be adequate, with the third of doubtful importance.

The factors seem to be open to about the same interpretation in this case as in the case of the 32 variables, except that the second and third factors have been interchanged. In this case, the second factor is the CR factor and the third factor the transfer factor, if we choose to adopt this interpretation of them.

The number of variables was further reduced to thirteen. In fitting a factor pattern to these, only two factors were fitted. As we shall see presently from a consideration of the residuals, two factors seemed ample

for this set of correlations. The factor loadings appear in Table 9.

TABLE 9

FACTOR LOADINGS OF 13 VARIABLES

Variable	Factor 1	Factor 2
1	.412	.061
2 etc., 6 etc.	.727	—.158
5, 9	.736	—.353
10 etc., 15 etc.	.606	—.161
13, 18	.597	—.444
20	.174	.179
22	.752	—.323
23	.773	—.025
24	.368	—.135
25, 27	.362	.707
26, 28	.501	.773
29, 30	.249	.047
32	.497	—.164
$\frac{\Sigma k^2}{n}$	.307	.127

In this case we have the residuals for only the first two factors:

	First residuals %	Second residuals %	Normal probability %
Within 1 sigma	65.4	85.9	68.3
2	93.6	97.5	95.4
3	98.8	100.0	99.7

In this case the first factor seems to be the factor of what is general to the various tests, highly weighted for the maze scores and for the time scores on the Jenkins problem box. The second factor gives a large positive weight to the CR scores, and negative weights to most of the rest of the scores. The scores having the largest negative weights are time scores. It is not surprising that we find no considerable transfer factor with this

group of variables, as we have combined the first and second patterns of the mazes and of the CR test into single scores.

At this point let us compare the factor pattern obtained for the 13-variable array by the Thurstone center-of-gravity method with the one obtained by Hotelling's method. The factor loadings for the two methods are given in Table 10.

TABLE 10

FACTOR LOADINGS OF 13 VARIABLES

Variable	Factor 1 Thurstone	Factor 1 Hotelling	Factor 2 Thurstone	Factor 2 Hotelling
1	.41	.49	.06	.21
2 etc., 6 etc.	.73	.74	—.16	.01
5, 9	.74	.80	—.35	—.26
10 etc., 15 etc.	.61	.60	—.16	—.01
13, 18	.60	.69	—.44	—.39
20	.17	.12	.18	.32
22	.75	.81	—.32	—.18
23	.77	.75	—.02	.07
24	.37	.41	—.13	—.08
25, 27	.36	.23	.71	.83
26, 28	.50	.39	.77	.80
29, 30	.25	.26	.05	.12
32	.50	.55	—.16	—.12

We see that the loadings differ somewhat in detail, but that the general pattern is about the same in both cases. This is seen if we arrange the variables in order of size of the loadings for a particular factor and compute a rank-difference correlation between the orders obtained by the two methods. The correlation is found to be over .95 for each factor. The loadings by Hotelling's method all tend to be shifted somewhat, but in most cases the shift is in the same direction for each variable. Using this method does not seem to change

or add to our essential results.

Perhaps a graphic presentation of the factors will show up certain relationships which the numerical results do not readily show. Where three factors have been fitted, a tri-dimensional figure would be required to show up the relationships. The factors can be shown in bi-dimensional form by taking them in pairs. We have done this for the case of 32 variables. In the 20-variable and 13-variable factor patterns, we have

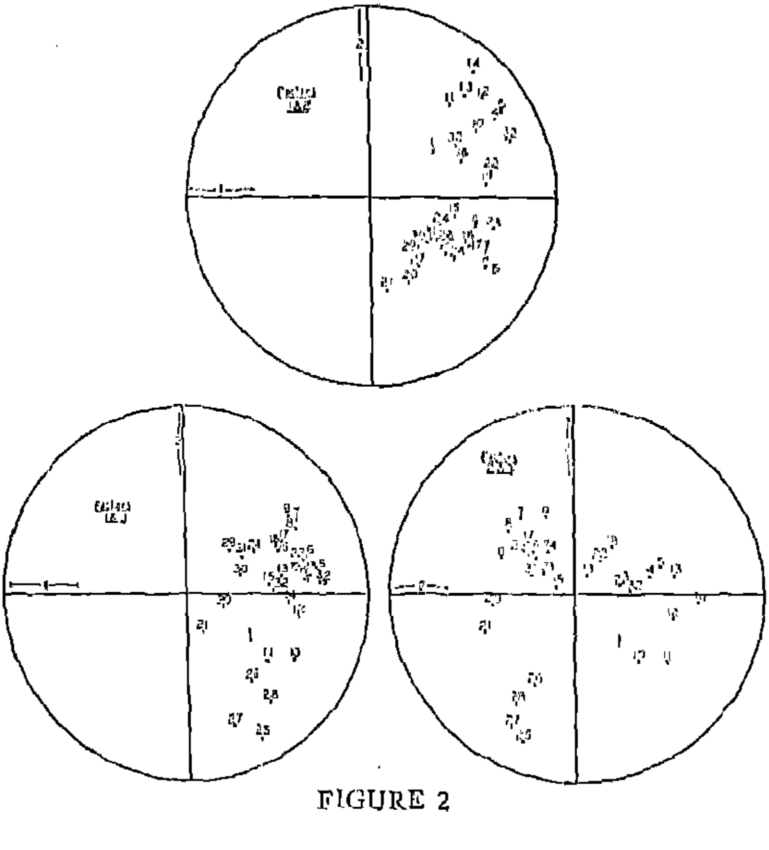


FIGURE 2

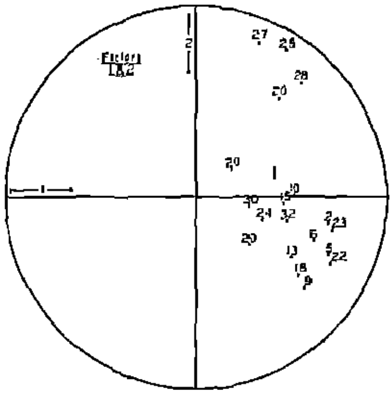


FIGURE 3

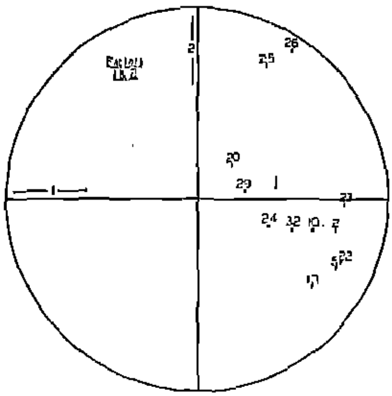


FIGURE 4

graphed only the first two factors. In the graphs, each variable is represented by a point. The identifying number of the variable is above the point, whereever possible.

When we examine the graph of the first two factors for the 32 variables, we notice first of all that the different measures taken on the same piece of apparatus tend

to cluster together. We find all the different scores for maze A, or for the CR apparatus, or for the Columbia obstruction apparatus in the same general locality. We can also see that the earlier tests tend to be in the upper half of the cluster. Finally, most of the right-hand half of the cluster is composed of maze and problem-box scores.

The striking characteristic of the clustering for factors one and three is the separation of the CR tests from the rest of the cluster.

The graph for factors one and two of the 20-variable array of correlations is very similar to that for factors one and three of the 32-variable problem just considered. The CR scores are isolated from the rest of the scores. It is interesting to note that the five scores furthest removed from the CR scores are time scores. In this case, the second factor appears to distinguish between rapid running of a maze or problem box and frequent crossings of the fence in the CR apparatus.

The 13-variable graph resembles that for 20 variables and shows the same trends.

We may make one more approach to the analysis of the intercorrelations by finding to what extent factors common to two different mazes and those common to two different problem boxes, for instance, are common between the mazes and problem boxes. We use the more general case of Spearman's attenuation formula, which takes the form

Community of a and b with x and y =

$$\frac{\sqrt[4]{r_{ax} \cdot r_{bx} \cdot r_{ay} \cdot r_{by}}}{\sqrt{r_{ab} \cdot r_{xy}}}$$

Comparing the two alley mazes with the two elevated mazes, we find a community of .98 for errors and .94 for time. Comparing scores on two different mazes with time scores on two different problem boxes, we get communities ranging from .70 to 1.00+, with a median at .91. On the other hand, when we take two maze or problem-box scores and get the community with two CR scores, the communities are very low. Four comparisons made at random give communities of .36, .11, .00, and .27. Likewise, when we consider the community between the two hunger-drive scores and a pair of maze, problem-box, or CR scores, we get sample results of .43, .27, .25, and .36.

This analysis suggests that the different mazes and problem boxes have very nearly the same factors in common, but that most of the factors common to the CR scores or to the hunger drive scores are common to those scores alone.

VI

DISCUSSION

Our results have shown a general tendency for different abilities in the same rat to be correlated. The relationship between different mazes is considerable—the median corrected correlation between scores on the different mazes that we used being about .45. But the relationships between the other tests are generally slight. The general tendency for correlation is confirmed in the work of Dunlap (5), and the higher relationship among mazes is in accord with the work of Commins, McNemar and Stone (3), and of Tomilin and Stone (33).

When we fitted factor patterns to our array of intercorrelations, using three factors, the first factor was weighted positively for each variable, and seemed to be composed of what was common to all the variables, especially what was common to the mazes and problem boxes. Another factor seemed to be a factor of transfer, or improved adjustment to the situation. The third factor seemed to be concerned chiefly with the CR tests.

The experimenter feels, from his observation of the rats at work, that a very important element in most of the time scores, and even in the error scores at the beginning of the test, was the "tameness" of the animals. It was obvious that some rats "knew" the maze long before they started to make perfect trials. The scores of the poor rats were probably poor largely because they were nervous in the maze and slightly conditioned against the food box. The scores of the good rats were

probably more nearly measures of their ability to learn the maze. Thus, the maze-learning score would not be a homogeneous function, but would measure "intelligence" for good rats and "tameness" for the poor ones.

Under these circumstances, it seems possible that the first factor is this combination of "intelligence" and "tameness" which goes to explain the scores. We might call it "docility." The "transfer" factor might be thought of as a change in the relative importance of tameness in the scores of certain rats. It might represent the decreasing importance of tameness and the increasing importance of ability in the later scores.

The CR factor seems to be concerned largely with these tests, and is probably due in large part to the common apparatus and procedure and to the contiguity in time of these tests. The high correlation between the scores on jumping the fence to the stimulus and jumping the fence between stimulations suggests that our training has merely increased the readiness of the response, instead of attaching it to a particular stimulus. What we have called "conditioned-response" scores may be measures of irritability or sensitivity to shock rather than of learning. In this connection, we observe an appreciable correlation between these scores and activity as measured by the revolving wheel.

In conclusion, it must be remembered that any number of factor patterns can be fitted to a set of correlations such as these, and that many of them will give about as satisfactory fits. Even if we admit these particular factors as the best ones, there may still be con-

siderable latitude in interpreting them. We must realize that the factors as they stand do not coincide with unitary biological traits, and at best only approximate such traits. We must look upon any names that we apply to individual factors purely as convenient tags, to be viewed with considerable distrust.

VII

SUMMARY AND CONCLUSIONS

1. Records were obtained for sixty-four albino rats on a wide variety of tests, including mazes, problem boxes, revolving-wheel activity cage, conditioned-response tests, and a measure of hunger drive.

2. The reliabilities of the various scores were determined, in most cases, from the scores on odd and even days corrected by the Brown-Spearman formula. With the exception of the scores for running the elevated mazes, most of these corrected reliabilities were fairly high (.70 to .95).

3. The intercorrelations of the different tests were computed. These were found to be positive in about 85 per cent of the cases. Most of the correlations between different tests were quite low, except for the correlations between different mazes. The median correlation between mazes was about .45 or .50 while the correlations between tests which were not both mazes had a median value of about .20.

4. A factor pattern was fitted to the intercorrelations by Thurstone's center-of-gravity method. More than one factor seemed necessary in order to get a satisfactory fit. The following factors were tentatively identified: (1) docility—maze-learning, intelligence, tameness; (2) transfer—distinguishing early from later tests; (3) a factor specific to the different CR scores.

References

1. Bagg, H. J. Individual differences and family resemblances in animal behavior. *Arch. Psychol.*, 1920, No. 43. Pp. 58.
2. Commins, W. D. A note on the learning of elevated mazes by rats. *J. Genet. Psychol.*, 1932, 41, 481-482.
3. Commins, W. D., McNemar, Q., & Stone, S. P. Intercorrelations of measures of ability. *J. Comp. Psychol.*, 1932, 14, 225-236.
4. Dennis, W. A block elevated maze for rats. *J. Comp. Psychol.*, 1931, 12, 429-432.
5. Dunlap, J. W. The organization of learning and other traits in chickens. *Comp. Psychol. Monog.*, 1933, 9, No. 44. Pp. 54.
6. Heron, W. T. The reliability of the inclined plane problem box as a measure of learning ability in the white rat. *Comp. Psychol. Monog.*, 1922, 1, No. 1, 1-36.
7. ———. The test-retest reliability of rat learning scores from the multiple-T maze. *J. Genet. Psychol.*, 1930, 38, 101-113.
8. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 1933, 24, 417-441, 498-520.
9. Hunter, W. S. Correlation studies with the maze in rats and humans. *Comp. Psychol. Monog.*, 1922, 1, No. 1, 37-56.
10. ———. Habit interference in the white rat and in human subjects. *J. Comp. Psychol.*, 1922, 2, 29-60.
11. Hunter, W. S., & Randolph, V. Further studies on the reliability of the maze with rats and humans. *J. Comp. Psychol.*, 1924, 4, 431-442.
12. ———. A note on the reliability of the maze as a method of learning in the Angora goat. *J. Genet. Psychol.*, 1926, 33, 3-8.
13. Husband, R. W. A comparison of human adults and white rats in maze learning. *J. Comp. Psychol.*, 1929, 9, 361-377.
14. Kelly, T. L. Statistical methods. New York: Macmillan, 1923. Pp. 300.
15. Leeper, R. The reliability and validity of maze experiments with white rats. *Genet. Psychol. Monog.*, 1932, 11, 141-173.

16. LIGGETT, J. R. A note on the reliability of the chick's performance in two simple mazes. *J. Genet. Psychol.*, 1925, 32, 470-480.

17. MILES, R. W. The comparative learning of rats on elevated and alley mazes of the same pattern. *J. Comp. Psychol.*, 1930, 10, 237-261.

18. RIESS, B. F. Limits of learning in the white rat and the guinea pig. *Genet. Psychol. Monog.*, 1934, 15, 305-368.

19. RUCH, F. L. Food-reward versus escape-from-water as conditions motivating learning in the white rat. *J. Genet. Psychol.*, 1930, 38, 127-145.

20. RUNDQUIST, E. A. Inheritance of spontaneous activity in rats. *J. Comp. Psychol.*, 1933, 16, 415-438.

21. SHIRLEY, M. Studies of activity. I. Consistency of the revolving drum method of measuring the activity of the rat. *J. Comp. Psychol.*, 1928, 8, 23-38.

22. ———. Studies in activity. IV. The relation of activity to learning and to brain weight. *J. Comp. Psychol.*, 1928, 8, 187-196.

23. SHUEY, A. M. The limits of learning ability in kittens. *Genet. Psychol. Monog.*, 1931, 10, No. 4, 287-378.

24. SPENCE, K. W. The reliability of the maze and methods of its determination. *Comp. Psychol. Monog.*, 1932, 8, No. 40. Pp. 45.

25. STONE, C. P. The reliability of rat learning scores obtained from a modified Carr maze. *J. Genet. Psychol.*, 1928, 35, 507-521.

26. ———. A multiple discrimination box and its use in studying the learning ability of rats. I. Reliability of scores. *J. Genet. Psychol.*, 1928, 35, 557-573.

27. STONE, C. P., & NYSWANDER, D. The reliability of rat learning scores from the multiple-T maze as determined by four different methods. *J. Genet. Psychol.*, 1927, 34, 497-524.

28. THURSTONE, L. L. The theory of multiple factors. Ann Arbor, Mich.: Edwards, 1933. Pp. 65.

29. ———. A simplified multiple-factor method. Ann Arbor, Mich.: Edwards, 1933. Pp. 25.

30. TOLMAN, E. C. The inheritance of maze-learning in rats. *J. Comp. Psychol.*, 1924, 4, 1-18.

31. TOLMAN, E. C., & DAVIS, F. C. A note on the correlation between two mazes. *J. Comp. Psychol.*, 1924, 4, 125-136.

32. TOLMAN, E. C., & NYSWANDER, D. The reliability and validity of maze-measures for rats. *J. Comp. Psychol.*, 1927, 7, 425-460.
33. TOMILIN, M. I., & STONE, C. P. Intercorrelation of measures of learning ability in the albino rat. *J. Comp. Psychol.*, 1934, 17, 73-88.
34. TRYON, R. C. Studies in individual differences in maze ability. I. The measurement of the reliability of individual differences. *J. Comp. Psychol.*, 1930, 11, 145-170.
35. ———. Studies in individual differences in maze ability. III. The community of function between two maze abilities. *J. Comp. Psychol.*, 1931, 12, 95-115.
36. WARDEN, C. J. Animal motivation. New York: Columbia Univ. Press, 1931. Pp. 502.
37. WARDEN, C. J., JENKINS, T. N., & WARNER, L. H. Introduction to comparative psychology. New York: Ronald, 1934. Pp. 581.
38. WARNER, L. H. Association span of the white rat. *J. Genet. Psychol.*, 1932, 41, 63-65.
39. WARNER, L. H., & WARDEN, C. J. The development of a standardized animal maze. *Arch. Psychol.*, 1927, No. 92. Pp. 35.
40. WEBB, L. W. Transfer of training and retroaction. A comparative study. *Psychol. Rev., Monog. Suppl.*, 1917, 24, No. 3. Pp. 90.
41. WILLIAMS, K. A. The reward value of a conditioned stimulus. *Univ. Calif. Publ. Psychol.*, 1929, 4, No. 3, 31-55.

L'ORGANISATION DU COMPORTEMENT CHEZ LE RAT BLANC

(Résumé)

On a testé soixante-quatre rats blancs, âgés de soixante jours, dans dix divers appareils. On a testé les animaux sur deux formes du labyrinthe à unités multiples de Warner-Warden, deux formes d'un labyrinthe élevé, deux types de boîte à problème, deux tests de réponse conditionnelle, l'appareil d'obstruction de Columbia avec une impulsion de faim de vingt-quatre heures, et la cage d'activité à roue tournante. Quelques-uns des tests ont donné plusieurs différents types de résultats, de sorte que l'on a obtenu des dix tests trente-deux résultats de temps, d'erreur, et de distance.

Dans la plupart des cas on a déterminé les constances des divers résultats au moyen des résultats les jours pairs et impairs corrigés par la formule de Brown-Spearman. A l'exception des résultats du parcours du labyrinthe élevé, la plupart de ces constances corrigées ont été assez élevées,

c'est-à-dire, de 0,70 à 0,95. Les résultats des tests n'ont pas été influencés d'une façon marquée par les variations quotidiennes du rendement.

On a computé les intercorrélations des divers tests. Celles-ci se sont montrées positives dans environ 85 pour cent des cas, et nulle des corrélations négatives n'a été constamment différente de nulle. La plupart des corrélations entre les divers tests ont été assez peu élevées, à l'exception des corrélations entre les divers labyrinthes. La corrélation médiane corrigée entre les labyrinthes a été d'environ 0,45 ou 0,50, tandis que les corrélations corrigées entre les tests qui n'ont pas été tous deux des labyrinthes ont eu une valeur médiane d'environ 0,20.

On a ajusté une forme de trois facteurs aux corrélations entre les trente-deux variables, et aussi à une série réduite de corrélations entre vingt variables, au moyen de la méthode du centre de gravité de Thurstone. Le premier facteur a eu un poids positif pour chacune des variables, et a semblé surtout associé aux labyrinthes et aux boîtes à problème. Le deuxième facteur a donné des poids positifs aux premiers tests et des poids négatifs aux subséquents. Le troisième facteur a donné de grands poids négatifs aux tests de réponse conditionnelle et de petits poids négatifs ou positifs à tous les autres. Les trois facteurs ont été provisionnellement identifiés comme (1) la capacité en train d'être testée—l'apprentissage du labyrinthe, "l'intelligence," "la docilité," (2) un facteur de "transfert," discriminant entre les premiers tests et les subséquents, et (3) un facteur qui a eu surtout rapport, paraît-il, aux tests de réponse conditionnelle.

THORNDIKE

DIE ORGANISATION DES VERHALTENS DER ALBINORATTE

(Referat)

Vierundsechzig Albinoratten im Alter von sechzig Tagen wurden in zehn verschiedenen Apparaten untersucht. Die Tiere wurden bei zwei Mustern des Warner-Warder vielfachen Einheitslabyrinths, zwei Mustern des erhobenen Labyrinths, zwei Arten des Problemkastens, zwei bedingten Antwortstests, dem Columbia Hindernisapparat nach vierundzwanzig Stunden Futterentziehung, und dem Drehradtätigkeitskäfig untersucht. Einige Versuche ergaben mehrere Typen von Ergebnissen. Auf dieser Weise wurden im ganzen zweiunddreissig Zeit- , Irrtum- , und Entfernungswerte von den zehn Versuchen gewonnen.

Die Zuverlässigkeiten der verschiedenen Resultate wurden bestimmt; in den meisten Fällen wurden Ergebnisse an geraden und ungeraden Tagen durch die Brown-Spearman Formel verbessert. Mit der Ausnahme der Ergebnisse für das erhobene Labyrinth waren die meisten dieser verbesserten Zuverlässigkeiten ziemlich hoch, d.h. 0,70 bis 0,95. Die Ergebnisse wurden durch tägliche Veränderungen der Leistung nicht ungebührlich beeinflusst.

Die Zwischenkorrelationen der verschiedenen Versuche wurden ausgerechnet. Es wurde festgestellt, dass diese in ungefähr 85 Prozent der Fälle positiv waren, und keine der negativen Korrelationen zuverlässig verschieden von Null war. Die meisten Korrelationen zwischen den verschiedenen Versuchen waren ganz niedrig, mit der Ausnahme der Korrelationen zwischen den verschiedenen Labyrinthen. Die verbesserte Durchschnittskorrelation zwischen den Labyrinthen war ungefähr 0,45 oder 0,50 während die verbesserte Korrelationen zwischen den Versuchen, die nicht von beiden Labyrinthen waren, einen Durchschnittswert von ungefähr 0,20 hatten.

Das Gebilde von drei Faktoren wurde den Korrelationen zwischen den

zweiunddreissig Variablen passend gemacht, und auch einer verminderten Reihe von Korrelationen zwischen zwanzig Variablen vermittelst der Thurstone Schwerpunktmethode (center-of-gravity method). Der erste Faktor hatte ein positives Gewicht für jede der Variablen und schien besonders mit den Labyrinthen und Problemkasten in Beziehung zu stehen. Der zweite Faktor ergab positive Gewichte bei den früheren Versuchen und negative Gewichte bei den späteren. Der dritte Faktor ergab grosse negative Gewichte bei den bedingten Antwortsversuchen und kleine negative oder positive Gewichte bei den anderen. Die drei Faktoren wurden probend identifiziert als (1) die untersuchte Fähigkeit—das Labyrinthlernen, "Intelligenz," "Fügsamkeit," (2) ein Faktor der "Übertragung," das Unterscheiden zwischen den früheren und späteren Versuchen, und (3) ein Faktor, der sich scheinbar besonders auf die bedingten Antwortsversuche bezog.

THORNDIKE

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

April, 1935
Volume XVII, No. 2

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at
Worcester, Mass., under Act of March 3, 1879

Journal of General Psychology

EDITED BY

CARL MURCHISON, *Clark University*

Frank Angell, *Stanford University*; F. C. Bartlett, *Cambridge University*; V. M. Borovski, *Gosudarstvennyi Institut Eksperimental'noi Psikhologii, Moskva*; G. S. Brett, *University of Toronto*; Karl Bühler, *Universität Wien*; Harvey Carr, *University of Chicago*; W. J. Crozier, *Harvard University*; Raymond Dodge, *Yale University*; James Drever, *University of Edinburgh*; Georges Dumas, *Université de Paris*; Knight Dunlap, *The Johns Hopkins University*; H. L. Hollingworth, *Columbia University*; Clark L. Hull, *Yale University*; T. A. Hunter, *Victoria University College, New Zealand*; Pierre Janet, *Université de Paris*; F. Kiesow, *Regia Università degli Studi, Torino*; Otto Klemm, *Universität, Leipzig*; Felix Krueger, *Universität, Leipzig*; Matataro Matsumoto, *Tokyo Imperial University*; William McDougall, *Duke University*; A. Michotte, *Université de Louvain*; Walter Miles, *Yale University*; John Paul Nafe, *Washington University*; Samuel Renshaw, *The Ohio State University*; Kanae Sakuma, *Kyushu Imperial University, Fukuoka*; A. L. Schniermann, *Bekhterev Reflexological Institute for Brain Research, Leningrad*; C. E. Spearman, *University of London*; Edward C. Tolman, *University of California*; John B. Watson, *New York City*; H. P. Weld, *Cornell University*; Raymond H. Wheeler, *University of Kansas*; R. W. Wilcocks, *Universiteit van Stellenbosch, South Africa*; Robert M. Yerkes, *Yale University*.

Quarterly, published in January, April, July, and October. Devoted primarily to experimental, theoretical, clinical, and historical psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to the general editorial office. All subscriptions and business communications should be sent directly to the Clark University Press. Beginning January, 1930, 1000 printed pages (two volumes) annually. Per annum $14.00; per volume $7.00; single numbers $4.00. Complete volumes from 1927 at $7.00 and $9.00 per volume, plus transportation.

The Pedagogical Seminary and

Journal of Genetic Psychology

EDITED BY

CARL MURCHISON, *Clark University*

John E. Anderson, *University of Minnesota*; Charlotte Bühler, *Universität Wien*; William H. Burnham, *Clark University*; Cyril Burt, *University of London*; Leonard Carmichael, *Brown University*; Ed. Claparède, *Université de Genève*; Edmund S. Conklin, *Indiana University*; Sante De Sanctis, *R. Università di Roma*; Arnold Gesell, *Yale University*; William Healy, *Judge Baker Foundation, Boston*; Leta S. Hollingworth, *Teachers College, Columbia University*; Buford Johnson, *The Johns Hopkins University*; Harold E. Jones, *University of California*; Truman L. Kelley, *Harvard University*; Yoshihide Kubo, *Hiroshima Normal College*; K. S. Lashley, *Harvard University*; A. R. Luria, *Akademiia Kommunisticheskogo Vospitaniia im. N. K. Krupskoi, Moskva*; Toshio Nogami, *Kyoto Imperial University*; Ivan P. Pavlov, *Gosudarstvennyi Institut Eksperimental'noi Meditsiny, Leningrad*; Henri Piéron, *Université de Paris*; William Stern, *Hamburgische Universität*; Calvin P. Stone, *Stanford University*; Lewis M. Terman, *Stanford University*; Godfrey Thomson, *University of Edinburgh*; E. L. Thorndike, *Teachers College, Columbia University*; C. J. Warden, *Columbia University*; John B. Watson, *New York City*; Helen Thompson Woolley, *Teachers College, Columbia University*.

Quarterly, published in March, June, September, and December. Devoted to child behavior, animal behavior, and comparative psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to the general editorial office. All subscriptions and business communications should be sent directly to the Clark University Press. Beginning January, 1930, [illegible] printed pages (two volumes) annually. Per annum $14.00; per volume $7.00; single numbers $4.00. Complete sets from 1891 at $7.00, $8.00, and $10.00 per volume, plus transportation.

CLARK UNIVERSITY PRESS
Worcester, Massachusetts
U. S. A.

Genetic Psychology Monographs

EDITED BY

CARL MURCHISON, *Clark University*

and an international cooperating board which is the same as that of the *Journal of Genetic Psychology*.

Bi-monthly. One volume per year. Child behavior, animal behavior, and comparative psychology. Each number a complete research. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to the general editorial office. All subscriptions and business communications should be sent directly to the Clark University Press. Approximately five hundred pages annually. Per annum $7.00; single numbers $2.00; double numbers $3.00. Complete sets from 1926 at $7.00 and $9.00 per volume, *plus transportation*.

VOLUME I—1926

1. Performance Tests for Children of Pre-School Age—R. STUTSMAN
2. An Experimental Study of the Eidetic Type—H. KLÜVER
3 & 4. A Study of Natio-Racial Mental Differences—N. D. M. HIRSCH
5. A Psychological Study of Juvenile Delinquency by Group Methods—J. W. BRIDGES AND K. M. B. BRIDGES
6. The Influence of Puberty Praecox on Mental Growth—A. GESELL

VOLUME II—1927

1 & 2. The Mind of a Gorilla—R. M. YERKES
3. The Rôle of Eye-Muscles and Mouth-Muscles in the Expression of the Emotions—K. DUNLAP
4. Family Similarities in Mental-Test Abilities—R. R. WILLOUGHBY
5. Coordination in the Locomotion of Infants—L. H. BURNSIDE
6. The Mind of a Gorilla: Part II. Mental Development—R. M. YERKES

VOLUME III—January-June, 1928

1. An Experimental Study of the Olfactory Sensitivity of the White Rat—J. R. LIGGETT
2. A Photographic Study of Eye Movements in Reading Formulae—M. A. TINKER
3. An Experimental Study of the East Kentucky Mountaineers—N. D. M. HIRSCH
4. Responses of Foetal Guinea Pigs Prematurely Delivered—G. T. AVERY
5. Objective Differentiation between Three Groups in Education (Teachers, Research Workers, and Administrators)—M. B. JENSEN
6. The Effect of Segregation on the Sex Behavior of the White Rat as Measured by the Obstruction Method—M. JENKINS

VOLUME IV—July-December, 1928

1. Observation and Training of the Fundamental Habits in Young Children—E. A. BOTT, W. E. BLATZ, N. CHANT, AND H. BOTT
2 & 3. Determination of a Content of the Course in Literature of a Suitable Difficulty for Junior and Senior High School Students—M. C. BURCH
4 & 5. Methods for Diagnosis and Treatment of Cases of Reading Disability—M. MONROE (*foreword by Samuel T. Orton*)
6. The Relative Effectiveness of Lecture and Individual Reading as Methods of College Teaching—E. B. GREENE

VOLUME V—January-June, 1929

1. The Age Factor in Animal Learning: I. Rats in the Problem Box and the Maze—C. P. STONE
2. The Effect of Delayed Incentive on the Hunger Drive in the White Rat—E. L. HAMILTON
3. Which Hand Is the Eye of the Blind?—J. M. SMITH

Genetic Psychology Monographs (*continued*)

4. The Limits of Learning Ability in Kittens—A. M. Shuey
5 & 6. The Effect of Habit Interference upon Performance in Maze Learning— O. W. Alm

VOLUME XI—January-June, 1932

1. General Factors in Transfer of Training in the White Rat—T. A. Jackson
2. The Effect of Color on Visual Apprehension and Perception—M. A. Tinker
3. The Reliability and Validity of Maze Experiments with White Rats—R. Leeper
4. A Critical Study of Two Lists of Best Books for Children—F. K. Shuttleworth
5 & 6. *Measuring Human Energy Cost in Industry: A General Guide to the Literature*—R. M. Page

VOLUME XII—July-December, 1932

1. Family Resemblances in Verbal and Numerical Abilities—H. D. Carter
2. The Development of Fine Prehension in Infancy—B. M. Castner
3 & 4. *The Growth of Adaptive Behavior in Infants: An Experimental Study at Seven Age Levels*—H. M. Richardson
5 & 6. Differential Reactions to Taste and Temperature Stimuli in Newborn Infants—K. Jensen

VOLUME XIII—January-June, 1933

1. A Critique of Sublimation in Males: A Study of Forty Superior Single Men—W. S. Taylor
2. A Study of the Nature, Measurement, and Determination of Hand Preference—H. L. Koch, *et al.*
3. The Growth and Decline of Intelligence—H. E. Jones and H. S. Conrad
4. The Relation between the Complexity of the Habit to Be Acquired and the Form of the Learning Curve in Young Children—M. L. Mattson
5. Eating Habits in Relation to Personality Development of Two- and Three-Year-Old Children: A Study of Sixty-nine Children in Two Nursery Schools—A. A. Eliot
6. Coordinating Mechanisms of the Spinal Cord—O. C. Ingebritsen

VOLUME XIV—July-December, 1933

1. Mental Growth during the First Three Years: A Developmental Study of Sixty-one Children by Repeated Tests—N. Bayley
2. A Study of Triplets: Including Theories of Their Possible Genetic Relationships—F. N. Anderson and N. V. Scheidemann
3. The Objective Measurement of Emotional Reactions—H. V. Gaskill
4. Development of Behavior in the Fetal Cat—J. D. Coronios
5. A Study of Certain Language Developments of Children in Grades Four to Twelve, Inclusive—L. L. LaBrant
6. The Effect of Early and Delayed Practice on Memory and Motor Performance Studied by the Method of Co-Twin Control—J. R. Hilgard

VOLUME XV—January-June, 1934

1. Studies in the Psychology of Tone and Music—P. R. Farnsworth
2. Motor Learning of Children in Equilibrium in Relation to Nutrition—E. L. Breeb
3. *Discrimination Limens of Pattern and Size in the Goldfish Carassius auratus*—J. B. Rowley
4. Limits of Learning in the White Rat and the Guinea Pig—B. F. Riess
5 & 6. The Limits of Learning Ability in Rhesus Monkeys—H. A. Field

VOLUME XVI—July-December, 1934

1. A Statistical Study of Ratings on the California Behavior Inventory for Nursery School Children—H. S. Conrad
2. An Eye-Movement Study of Objective Examination Questions—A. Frandsen
3. An Experimental Study of Constitutional Types—O. Klineberg, S. E. Asch, and H. Block
4. The Development of a Battery of Objective Group Tests of Manual Laterality with the Results of their Application to 1300 Children—W. N. Durost
5 & 6. An Experimental Study in the Prenatal Guinea-Pig of the Origin and Development of Reflexes and Patterns of Behavior in Relation to the Stimulation of Specific Receptor Areas during the Period of Active Fetal Life—Leonard Carmichael

VOLUME XVII—January-December, 1935

1. Organization of Behavior in the Albino Rat—Robert Ladd Thorndike
2. Brightness Discrimination in the Rhesus Monkey—Meredith Pullen Crawford

---

CLARK UNIVERSITY PRESS
Worcester, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

April, 1935
Volume XVII, No. 2

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

BRIGHTNESS DISCRIMINATION IN THE RHESUS MONKEY

From the Animal Laboratory of the Department of Psychology, Columbia University

By

MEREDITH PULLEN CRAWFORD

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, in the Faculty of Philosophy, Columbia University.

NEW YORK
JANUARY, 1935

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

ACKNOWLEDGMENTS

The writer wishes to express his sincere appreciation to Professor C. J. Warden of the Department of Psychology of Columbia University for his advice throughout the experiment and for his thoughtful criticism in the preparation of the manuscript. Thanks also are due to Professor Henry W. Nissen, Dr. G. R. Wendt, and Mr. John T. Cowles of Yale University for their criticisms of the manuscript. We are particularly grateful to Mr. Richard H. Henneman, Mr. Robert I. Watson, and Miss Mary Waits who served as human subjects in this experiment.

MEREDITH PULLEN CRAWFORD

LABORATORIES OF COMPARATIVE PSYCHOBIOLOGY
YALE UNIVERSITY SCHOOL OF MEDICINE
NEW HAVEN, CONNECTICUT

CONTENTS

I

INTRODUCTION

The accurate determination of sensory limens in animal subjects is interesting from three points of view. First, with a knowledge of the limits of sensory acuity of the subject the investigator does not make the error of setting problems before the organism which are insoluble simply because *it does not have the sensory* capacity to respond to the critical stimuli. Also, the results of former investigations may be more correctly evaluated in the light of accurate information regarding the sensory capacity of the animal used. Second, while the comparative anatomy of receptors is often cited as evidence of similarity of sensory capacity, this evidence is not nearly as conclusive as that from studies of comparative behavior. Limen determinations made under similar conditions offer perhaps the best basis of comparison of the receptive capacities of different forms. Third, limen determinations with animals make *possible the verification of certain general principles* of the action of receptors which have been established mainly for man. From these three points of view probably no field offers problems more interesting than that of vision.

Even in the most casual observation one is impressed with the apparent keenness of vision of the monkey. His rapid shift of gaze, his close visual examination of new objects, and his deft manipulation of himself and parts of his environment suggest that he must see things clearly. His arboreal habitat in the wild and

his agile play on bars and swings in captivity seem to require that his vision be accurate and effective. That this is the case has long been the assumption of those interested in primate behavior.

Morphological evidence gives further support to this assumption. Binocular vision is an indication of the usefulness of the eye mechanisms. It has often been pointed out that in the primates there is an increase in the neural tissue associated with the visual function at the expense of that associated with olfaction. The anatomy of the monkey eye is similar to that of man and other primates. [See Hartman and Strauss (6) for an anatomical description of the eye of the Rhesus monkey.] In a study of the differentiation of the retina in various primates Wollard (19) made sections from three areas of the retinas of different primates. By comparison of the number of conducting segments, the degree of intermixture of rods and cones, and the extent and perfection of the fovea, it was not possible to differentiate the following primates with any degree of validity: man, chimpanzee, *Cercocebus fulginosus, Cebus fatuellus,* marmoset, and the Rhesus monkey. On the basis of retinal structure, then, we might expect the vision of these primates to be about equally good.

Behavioral evidence obtained in the laboratory is also indicative of well-developed vision in monkeys. Johnson (10) has studied the visual acuity of the Cebus monkey and found that it could discriminate white and black striae when they were in the ratio of 1:1.02 in width at two absolute levels of width. Klüver (12) has

performed tests of visual acuity on a number of different kinds of monkeys, and though no limens were determined keenness of vision was indicated in all subjects. Form vision has been studied by a number of different investigators including Bierens de Haan, Klüver (12), Watson (18), and most recently by Neet (12). All these studies have demonstrated the ability of the monkey to respond to small differences in size and shape of visual stimuli almost as well as man. Although color vision has been studied by Hess, and under better experimental conditions by Watson (18), the results so far indicate only that the monkey does respond to color differences; no chromatic limens have been determined. For a more extensive treatment of vision in monkeys and in primates in general see Warden, Jenkins, and Warner (16).

While more or less satisfactory work has been done on other forms of vision in the monkey, brightness discrimination has received little careful experimental study. Kinnaman (11) trained two young Rhesus monkeys to make a differential response to food containers around which pieces of grey papers of known brightness were placed. His results indicated that a difference of 9 per cent between the brightness of the papers was necessary for discrimination. Haggerty's (5) investigation gives further indication that the monkey distinguishes differences in brightness, though no limens are presented. Shepherd (14) used black and white cards as differential stimuli with monkeys, and Klüver (12) has employed varying shades of grey

papers of known brightness as stimuli in choice responses made by monkeys with the pulling-in technique. Up to the present time, however, there has been no study in which the ability of the monkey to discriminate differences in brightness has been systematically explored over one or more absolute levels and limens determined.

The present experiment was planned to make such a study of the brightness sensitivity of the Rhesus monkey (*Macacus mulattus*). In carrying out this program it seemed best to determine discrimination limens at several levels over the middle range of daytime vision with the light-adapted eye. It was hoped in this experiment to see if the brightness-sensitivity curve for the monkey is similar to that of man, with particular reference to the manner in which the relative difference threshold varies over the range of brightness used. As seen from the above discussion of the literature, this represents the first attempt at a study of this type in monkeys.

II

METHOD AND PROCEDURE

The present experiment was carried on in the Laboratory of Comparative Psychology, Columbia University, and was in progress from December 15, 1933, until August 15, 1934. Four Rhesus monkeys were trained in a discrimination set-up to reach for food toward the brighter window of a Yerkes-Watson light-discrimination apparatus. Limens were determined from the composite results of a number of experimental periods during which the subjects were required to judge stimuli of progressively decreasing brightness difference. Experimentation was done in a dark room which was especially designed for testing sensory capacities in animals. It was located on the quiet or campus side of the building and was fitted with double doors leading into the hallway and lightproof doors in front of the window. These lessened the noises coming from the outside. The walls were treated with a sound-absorbing material which served to deaden the noises made inside. The experimental room was connected with the vivarium by a public hallway through which the animals were brought in a rubber-tired truck to and from the apparatus. The animal subjects were housed in the primate section of the vivarium and were kept under uniform conditions throughout the experiment.

It will be convenient to discuss the method and procedure of this experiment under the following headings:

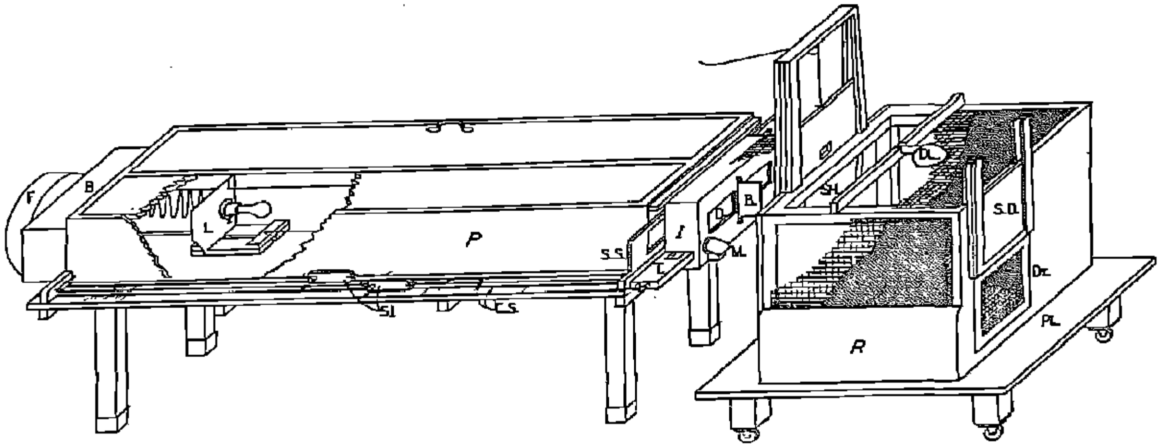


FIGURE 1

GENERAL VIEW OF THE APPARATUS AS ARRANGED FOR THE MONKEYS

The control-table is not shown. The three general divisions of the apparatus are: *P*, the modified form of the Yerkes Watson photometer box: *I*, the incentive-box; and *R*, the reaction-cage. The reaction-cage is moved further away from the incentive box than when in use.
Parts of each of the three general divisions are: Photometer; *B*, blower chamber with electric fan guard, *F*, attached to rear (electric fan used as a blower not shown); *L*, lamp carriage connected by cords over pulleys to slides, *Sl*, operating in grooves along the calibrated scale, *C.S.* The stimulus shifter, *S.S.*, is shown between the photometer and the incentive box.
Incentive-box: the food tray, *T*, which is open, the glass doors, *D*, the locking magnet, *M*, the blind, *B*, separating the space in front of the glass doors.
The reaction-cage, *R*, is shown mounted on a rolling platform *PL*. In the front wall is the shutter opening, *SH*, and in the rear a door opening outward, *Dr*, to which is attached a smaller vertically sliding door *S.D.* *DL* is the dome light over the reaction-cage. Screening is shown over ½-in. wire netting.

a. Apparatus.
b. Subjects.
c. Preliminary procedure.
d. Procedure of the experiment proper.

Apparatus

The apparatus used in this experiment was constructed in four parts, the photometer, the incentive-box, the reaction-cage, and the control-table. In Figure 1 three of these parts are represented together in

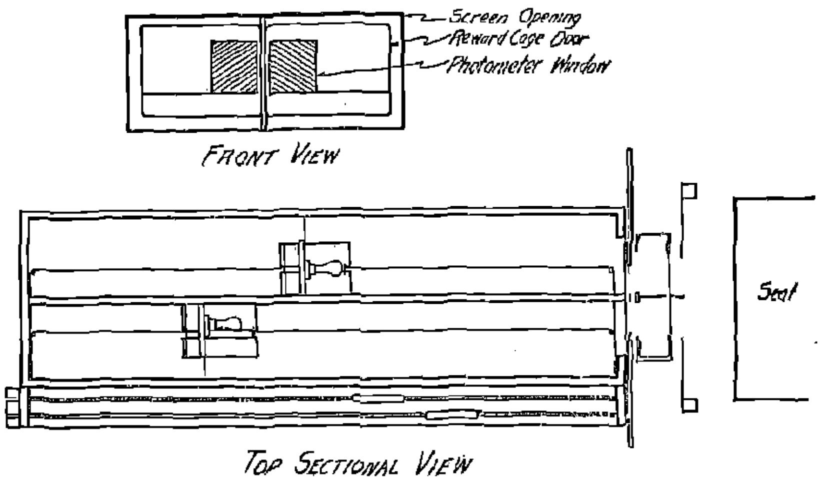


FIGURE 2

Sectional Views of the Apparatus as Arranged for Human Subjects

A. Top sectional view showing photometer box *P*, incentive box *I*, screen, *Sc*, and seat for observer. The legend is as in Figure 1 above.

B. Schematic view looking through opening in screen (or through shutter opening of animal reaction-cage), through the glass doors of the incentive-box (Reward-Cage), to the stimulus patches at the end of the photometer. Strips of paper are indicated as pasted over the lower edges of the glass doors of the incentive-box.

isometric projection, the control-table being omitted. Figure 2 gives a diagrammatic view of the apparatus,

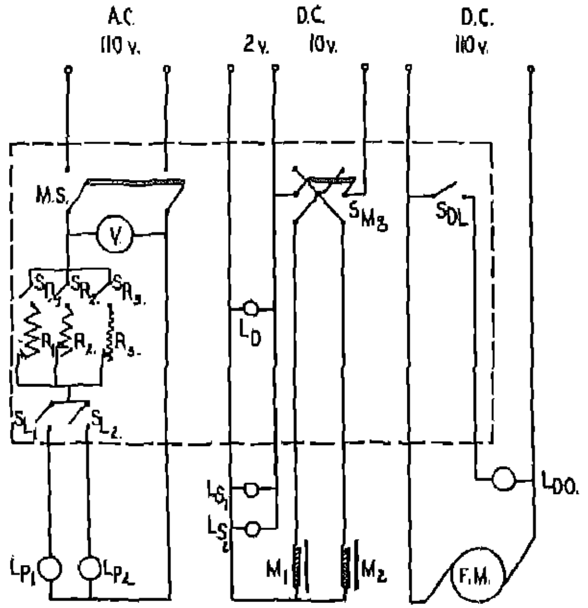


FIGURE 3

SCHEMATIC WIRING DIAGRAM OF THE APPARATUS

The devices on the control-table are enclosed within the broken line. The three sources of power are represented at the top, 110 v. A.C., 2-10 v. D.C., and 110 v. D.C. The legend is as follows.

M.S. master switch in 110 v. A.C. line.

$S_{R_{1,\,2\ and\ 3}}$ switches for operation of resistances.

$R_{1\ 2\ and\ 3}$ resistances for use *with different bulbs in* photometer.

$L_{p_{1\ and\ 2}}$ photometer bulbs in light carriages.

V voltmeter.

S_{MS} double pole double throw magnet switch.

$M_{1\ and\ 2}$ magnets on incentive box.

L_D 2-v. lamp on control-desk.

$L_{S_{1\ and\ 2}}$ 2 v. lamps on photometer slides.

S_{DL} switch operating dome light.

L_{DO} dome light in reaction-cage or over human screen.

F.M. blower fan motor.

exclusive of the control-table, as modified for use with human subjects. In Figure 3 a schematic wiring diagram of the electrical appliances used in the experiment is represented. The four main divisions of the apparatus, its modifications for use with human subjects, and the calibration of the photometer and illumination devices are described in separate sections below.

The Photometer. For quantitative measurement and presentation of the visual stimuli the light-box designed by Yerkes and Watson (20) was used. The apparatus will be referred to hereafter as the photometer. Certain minor modifications of the light carriages and the mechanism for moving these from the outside were made, though the essential nature of the device was preserved in the form described by the designers. The work of Graham and Nafe (4) with human subjects on intensity discrimination with the Yerkes-Watson apparatus has indicated that it is not a very reliable nor a very accurate instrument for studying brightness vision. The absolute limens which they determined were high and the variability was great. As shown by the results we obtained with human subjects (see pp. 142-146), the limens we determined were a good deal lower and more consistent over the levels tested. Our experimental technique differed from that of the above investigators, a fact which may in part account for the difference in results. The absolute values and the consistency of the limens which we obtained with human subjects make us feel justified in using the apparatus. It is conceivable, of

course, that another sort of light apparatus might have yielded slightly lower limens for both animal and human subjects.

The photometer consisted in an oblong wooden box 87 ins. long, 24 ins. wide, and 12 ins. high, heavily constructed of 1¼-in. pine stock. It was divided lengthwise by a partition ⅞ in. thick down the center, and was covered by two wooden hinged tops. Round openings, 4 ins. in diameter, were made in the rear end of each side of the photometer and were covered by brass Aubert diaphragms. The front end of the photometer was closed save for a horizontal opening running the entire width, 5¼ ins. in height Outside this horizontal opening were fitted brass guides into which two pieces of flashed opal glass (3-16 ins. thick) were placed one in front of each side of the photometer. A strip of wood separated the ends of these pieces of glass from each other. The whole photometer was mounted on a standard laboratory table 30 ins. high.

The lamp carriages were made of wood so treated as to slide noiselessly in the photometer. The base of each consisted of two parallel strips 10 ins. long and 3 ins. wide and separated by a space ½ in. wide. The strips fitted on either side of an iron rail which served as a guide along the floor of each side of the photometer. On this double base was mounted a vertical piece carrying a pasteboard face covering the entire cross-section of one side of the photometer. On this face was mounted a standard-size incandescent bulb socket (for a discussion of bulbs used see the section on

calibration below), so fitted as to be capable of motion in the horizontal plane. It was located at such a height that the horizontal axis of a bulb placed therein was always in a plane bisecting the opening in the front of the photometer covered by the opal glass. The electric lamp cords attached to the sockets of the carriages were fitted with rings which were strung on brass guide wires, one on each side of the photometer. Movement of the carriages moved the ring slides and thus took up the slack in the lamp cord.

In order to move the lamp carriages from outside the photometer a system of cords and pulleys was arranged. To one side of the photometer, on the outside, were fastened three wooden strips running the entire length of the photometer, one 1 in. wide and the other two 2 ins. wide. They were separated by two spaces, each 1/8 in. wide. Into these two spaces were fitted two wooden slides with bakelite bases, arranged so as to move the length of the photometer in these tracks. A brass edge was attached to each slide projecting out over the 2-ins. strips. Two-volt lamps were fitted to each slide and covered with small metal cylinders allowing the escape of only a small beam of light on each of the strips. By means of cords running over pulleys each of the slides was connected with one of the light carriages. Wire rope was first used over the pulleys, though since it was found that this frayed quickly, a good grade of manila cord, showing little tendency toward stretching, was finally adopted. With occasional replacements due to wear this cord proved satisfactory. The wooden strips

between which the slides moved were covered with paper on which scales of the different brightness levels were laid off.

At the rear end of the photometer was fitted a paste-board box to which the frame of a large Westinghouse electric fan was attached. Holes were bored in the sides of the photometer near the forward end, and were covered by hoods to prevent the escape of light. With the Aubert diaphragms open at the rear end of the photometer, and the fan in operation, a draught of air was forced through each side of the photometer, past the light carriages and bulbs, and out the ventilator holes in the sides. This forced draught ventilation system prevented overheating of the bulbs. The fan in operation also served as an effective sound screen.

In front of the flashed opal glass plates covering the forward openings of the photometer was fastened a steel rack. In this rack was placed an aluminum stimulus shifter, free to move back and forth across the face of the photometer. The shifter was fitted with two acid-blackened brass plates with square openings at their centers 2⅛ ins. on a side. The whole unit was similar to the stimulus adapter described by Yerkes and Watson. The shifter remained throughout the experiment in a constant position so that the brass plates outlined the stimulus patches of flashed opal glass. The patches were 1 in. apart. Since the subject's eye was from 9 to 11 ins. from the stimuli during the experiment, the stimulus patches subtended a visual angle of 5 or 6 degrees. The photometer was treated inside with several coats of dull black stain.

The Incentive-Box. The incentive-box was a device designed by the writer for presentation of the reward in close proximity to the stimuli. It might be used in manual discrimination set-ups with various sorts of visual stimuli. It was an oblong black box, 17 ins. wide, 4¾ ins. long, and 10¾ ins. high. A partition ¼ in. thick in the center of the box divided it into two compartments each 8⅜ ins. wide and 4¾ ins. long. The partition did not extend all the way to the floor of the box since clearance was left for the sliding tray which moved along the bottom of the box. In this tray were made two wells spaced so that when the tray was pushed into the incentive-box one well would be in each compartment. With the tray in this position no light could pass from one compartment to the other.

The incentive-box was mounted directly in front of the stimulus adapter on the end of the photometer. It was easily detachable therefrom by the removal of two bolts. The central partition of the incentive-box was in the same vertical plane as that of the photometer. In the front and back of the incentive-box two openings were made, 5½ ins. x 3¼ ins. directly in front of the stimulus patches defined by the brass plates in the stimulus shifter. The rear pair of openings in the incentive-box were left open, while the front pair were fitted with clear plate-glass doors, 7 x 5 3-16 ins. in dimensions, ⅜ in. thick, which overlapped the edge of the opening. They were made of *Vita Glass,* a product of the Vita Glass Corporation of New York. It is represented to have a transmission factor of .92 over

the visible range, and almost as high in the ultra-violet. The doors were hinged at the top and opened inward as viewed from the front of the incentive-box.

To the outer edge of each of these glass doors was bolted a strip of soft iron, 4 ins. x ½ in. x ⅛ in., projecting ¾ in. below the lower edge of the door. On the front face of the incentive-box, in the lower corners on either side, were mounted two electromagnets, each 1¼ ins. in diameter, and with a curved axis making an horizontal length of 3¼ ins. The soft iron laminated cores of these magnets projected through holes in the face of the incentive-box and came into contact with the free ends of the soft iron strips attached to the glass doors. When these magnets were supplied with sufficient current, they either held shut or firmly locked the glass doors, depending on the strength of the current supplied.

Thus in front view the box showed two identical glass doors, each directly in front of a stimulus patch on the end of the photometer. Through the door access to the food tray just below was gained. When one door was shut and the other locked with the differential electric current supplied, no visible cue was offered the observer indicating which door was shut and which one locked. This view is represented schematically in Figure 2 B.

On the narrow strip separating the two glass doors on the face of the incentive-box was mounted a wooden blind, 6 ins. high, ⅜ in. wide, and projecting out from the incentive-box 3⅛ ins. This blind projected the

same distance from the incentive-box as did the two electromagnets and their wooden bases. Around the front part of the incentive-box, attached to the sides, top, and bottom, and projecting out therefrom, was a piece of ½-in. 16-gauge wire netting. This was open only at the place where the food tray slid out of the incentive-box. The whole box was painted a dull black inside and out. Strips of black velvet cloth and cardboard were placed at points where the light was reflected from the stimulus patches. Two strips of black paper were placed over the lower part of the glass doors to prevent light from falling on the ledge just in front of the door, and to prevent the animal subject's direct vision of the food in the tray.

The Reaction-Cage. The platform which supported the reaction-cage was placed directly in front of the incentive-box. In Figure 1 the reaction cage is drawn away from the incentive-box in order to show the latter more clearly. During the experimentation the ends of the magnets on the incentive-box were almost touching the forward end of the reaction-cage. The movable platform on which the reaction-cage rested was 13½ ins. high and its top was 45 ins. long and 30 ins. wide. Four casters with 2-in. wheels were mounted at the bottom of the legs. The platform was satisfactorily solid even though mounted on casters.

The reaction-cage was built on a framework 2 ins. x 2 ins. square, and had dimensions of 30½ ins. x 27¼ ins. and 30¼ ins. At one end was placed a wooden shutter mechanism, consisting of an opening in the for-

ward cage wall 11¼ ins. x 5½ ins., through which the animal could see the stimulus patches and reach through the glass doors to the food tray in the incentive-box, and a vertically sliding shutter operated by a cord running over a pulley above. A small window in the shutter, 2 ins. x 1 in., allowed the animal to look at the stimulus patches but not reach for the doors when the shutter was raised halfway. Access to the cage was gained through a door in the other end, extending the full height of the cage and opening outward. On this door was fitted a smaller vertically sliding door to facilitate the transfer of animals from the transfer truck directly into the reaction cage.

On the roof of the cage ran a transverse piece to which was attached an electric bulb standard-size socket and a 6-in. desk-lamp reflector, called hereafter the dome light. The illumination furnished by the dome light in the reaction cage will be discussed below in the section on calibration. There was no floor in the reaction-cage since it rested directly, and without fastening, on the platform. There was an auxiliary platform of adjustable height which was placed in the reaction cage on the platform top for use with small animals, enabling them to reach the shutter opening when in a standing position. The reaction-cage was covered with ⅞-in. planking except for a part of one side, the top, the doorway, and shutter opening. A half-inch wire netting, 16-gauge, covered the side, top, and lower part of the door. Window screening, painted white, covered the wire netting. When the experimen-

tal room was darkened and the dome light turned on this screening acted as a one-way light screen. The inside of the reaction-cage, top of the moveable platform, and adjustable auxiliary platform were painted flat white, and the outside a darker color.

The Control-Table. The control-table was placed beside the forward end of the photometer scale and near the incentive-box. Its location was such that all the routine adjustments of the apparatus during experimentation, including observation of the subjects' responses, could be made with the minimum of movement by the experimenter. The table was 30 ins. high and had a top 33¾ x 21 ins. On the control table were located electric switches, resistances, a voltmeter, and a small 2-v. lamp socket to give light for recording. The various electrical appliances used in the apparatus were all operated from the control table. Each was connected in one of three circuits, of different power supply, as indicated in the wiring diagram given in Figure 3. The devices both on and off the control-table are represented in that figure. It will be convenient to indicate the connections of the three circuits separately below.

1. The 110-v. A. C. Circuit. The bulbs on the light carriages of the photometer were supplied with A. C. having a voltage fluctuation of 2-3 v. A master switch (*M.S.*, Figure 3), a voltmeter (V), three variable resistances (R_1, R_2, R_3), and switches for each (S_{R_1}, S_{R_2}, S_{R_3}), and switches for each photometer lamp (S_{L_1}, S_{L_2}) were located in convenient places on the control-table.

2. The 110-v D. C. Circuit. The bulb in the dome

light of the reaction-cage (or over the screen used with human subjects), and the electric fan were operated on 110-v D.C., which showed a fluctuation of 3-4 v. A switch (S_{DL}) operating the light was located on the control-table.

3. The 2-10-v. D. C. Circuit. The locking magnets on the incentive-box, the lights on the photometer slide, and the control-table light were operated on D.C. obtained from storage batteries. 2 and 10 volts were supplied to the arms of a double pole, double throw switch (S_{MG}) whose opposite binding posts were connected to the two magnets on the incentive-box (M_1, M_2). In one position of S_{MG}, M_1 would have a 2-v. potential impressed upon it and M_2 a 10-v. potential, and in the opposite direction of the switch the potentials in the two magnets would be reversed. This served to close one door and to lock the other with a single movement. When the magnets were in operation (i.e., during a discrimination response), the lights on the photometer slides (L_{B_1}, L_{B_2}) and the desk light (L_D) in the 2-v. circuit in parallel were made very dim.

All electric lines to and from the control-table were fitted with plugs and sockets so that the control-table could be easily disconnected from the rest of the apparatus. Beside the control devices permanently attached to the control-table, a stopwatch, container for food reward, and record blanks were on the table during experimentation.

Modifications of the Apparatus for Use with Human Subjects. For use with human subjects the reaction-cage was replaced by a white screen made of wallboard fastened to a wooden frame. The top of the screen

stood 53 ins. and the bottom 27 ins. from the floor. At the center of the screen was an opening, 8½ x 4 ins., covered by a shutter, operated up and down by a cord and pulley. The screen was placed directly in front of the incentive-box. When the shutter was up a view was afforded through the screen opening and incentive-box glass doors to the stimulus patches. This view is represented in Figure 2 B. From the usual position of the human subjects' heads the stimulus patches subtended a visual angle of about 5°.

A lampshade was placed above and in front of the screen in relatively the same position with respect to the shutter opening in the screen as was the dome light in the reaction-cage to the shutter opening there. A seat in front of the screen was so arranged for the human subjects, the small 2-v desk lamp and the shielded the stimuli through the shutter opening, were in the same relative position as that of the animal subject. A burlap screen was placed at the subjects' left, between them and the experimenter at the control-table, and a wall of the room was at the right. Thus, in effect, the enclosure was similar to that within the animal reaction-cage.

As noted above, the whole apparatus was used in a specially constructed dark-room, which was treated for sound deadening, though not soundproof. It was darkened during the experimental period except for such light as came from the dome light of the reaction cage or the light in front of the screen used with human subjects that their head and eyes, when regarding

photometer slide lamps, and, of course, that which came from the stimulus patches at the end of the photometer as it shone through the incentive-box.

Calibration of the Apparatus. The Macbeth Illuminometer was used for measurement of brightness in the apparatus. For calibration of the photometer a wooden crosspiece was mounted on the end of the blind on the incentive-box. This carried a notch at either end into which the object end of the illuminometer was fitted in such a manner as to point toward the stimulus patches. It was thus insured that the illuminometer when fitted into the notches always pointed in a constant direction, and that this direction corresponded to a subject's line of vision when regarding the stimuli. The measurements were made at about the distance from which the subject regarded the stimuli. It is known, however, that slight deviations from a constant direction of pointing of the illuminometer would not affect the brightness readings.

For any determination of brightness of the stimulus patches six or more readings were taken on each patch, the observations being made alternatively on both sides, and the settings of the illuminometer scale being made in ascending and descending directions in irregular order. The experimenter had had considerable practice in using the illuminometer before the final calibration readings were made. During measurement of the stimulus patches constant reference was made to the voltmeter in the photometer circuit, so that any fluctuations were quickly compensated for by a change in the proper resistance. The A. C. line supplying

the photometer bulbs was found to fluctuate two or three volts in either direction. Occasionally this change was sudden.

When calibration of the photometer was begun it was observed that the brightness over the whole area of each of the stimulus patches was uniform in so far as could be observed with the illuminometer. The characteristic curve of the whole photometer was first determined, using 40-w. bulbs in the lamp carriages operated at 106 v. It was found that this curve was far from what would be expected if the photometer followed the inverse square law. Modifications, such as a reduction in the amount of reflection from the inner walls of the photometer, were made and the characteristic curve redetermined. This time it was found more closely to approximate the inverse square relationship, though by no means exactly. Therefore, the assumption of the inverse square relationship over even a limited part of the photometer was considered unwise. Since this was the case, together with the fact that different-size bulbs were used at the different levels, it was necessary to calibrate the photometer for each brightness level separately. The approximate value of the standard brightness at each level was decided upon and then the photomoter calibrated around standard brightnesses as near these as was convenient.

Three sets of bulbs were used for the four levels. These appear in Table 1 below. It will be observed that the bulbs were operated at voltages below their normal ratings. This was done in order to decrease the change in their brightness due to age. As noted

TABLE I

BULBS AND THEIR OPERATING VOLTAGES USED AT EACH BRIGHTNESS LEVEL

Brightness level	Standard brightness	Bulbs used	Operating voltage*
I	0.0864 millilamberts	7-w. Westinghouse internal frosting.	111
II	0.7769 millilamberts	40-w. G.E. Mazda internal frosting.	106
III	7.146 millilamberts	40-w. G.E. Mazda internal frosting.	106
IV	55.3 millilamberts	300-w. G.E. Mazda clear.	94

*The operating voltage was subject to increase when it was observed that the bulb had become dimmer with use.

the voltage was increased when the bulbs showed signs of becoming dimmer. The particular voltages were chosen in order to give a color match at the stimulus patches with the standard field color of the Macbeth Illuminometer. This was a yellowish white and was uniform over the four brightness levels. Replacement of a pair of bulbs because of deviation from the adopted standard was necessary only in the case of the 40-w. pair used at levels II and III. A pair of bulbs was accurately matched before use in either side of the photometer.

The following procedure was adopted in the calibration for each brightness level. Positions of the lamp carriages were first determined at which the standard brightness was given at the stimulus patches. The pulley cords were then adjusted so that the slides were at corresponding (opposite) points on their respective scales. A rough determination was then made to locate the scale position of what probably

would be the brightest step at that level. The distance between the standard and this brightest position on the scale was divided into equal linear units and determinations of brightnesses made with each of the slides at each of these division points. Five or six such determinations were made on separate days and the brightness of each step was expressed as a multiple of the standard brightness as determined that day. Averages of these determinations were then made and a calibration curve for the photometer over that limited range and with that size bulb was drawn, relating brightness increment to linear distance on the scales.

From this curve points were laid off on the scale to indicate slide positions at which, when one slide was at the standard position, and the other at one of these positions, various steps of brightness difference would be presented between the two stimulus patches. Obviously, the patch on either side of the photometer might be made brighter by reversing the positions of the slides. The steps of difference in brightness used at the four levels were chosen as indicated in the chapter on results, the particular range being chosen after preliminary rough determination of the magnitude of the subjects' limens at that level. When calibration of the photometer was complete, the scales each bore four sets of markings, one set for each brightness level.

It was also necessary to measure the average brightness of the light in the animal reaction-cage and on and about the screen used for human subjects. This was done in a manner similar to the calibration of the photometer. "Illumination maps" of the screen and

cage were made showing the distribution of brightnesses over the various parts of the cage and screen. It was found that when a 25-w. internally frosted G.E. Mazda bulb was placed in the dome light of the reaction cage an average brightness of 1-12 millilamberts was present over the forward wall about the position of the shutter opening. On the side wall to the left the average brightness was only 1-3 millilamberts. This was caused by the presence of the screen through which a good deal of light passed. On the wall to the right the illumination was about as on the front wall, falling off toward an average value of 2-3 millilamberts toward the rear. On the rear wall the range was from 1 near the floor to 8 at the maximum, about two-thirds the way up. Thus the average brightness about the walls of the cage was in the region of 4-6 millilamberts. The floor, being of a darker color because of wear, measured 1-3 millilamberts. Since the animal was free to move about in the reaction-cage, and could easily look directly at the dome light, the possible range of the brightnesses to which he might have been exposed was a good deal greater than indicated by the brightness of the walls alone. The light when observed at a distance of two inches was as bright as the standard of level IV or brighter, being over 50 millilamberts. The average level of brightness over the whole volume of the reaction cage, then, would be difficult to determine. At least we may say that the range was from 1 to more than 50 millilamberts. The illumination conditions on and around the screen used for human subjects were essentially similar to those in

the animal reaction-cage when the same bulb was used in the reflectors of the dome light and the screen light. Also, the human subjects were free to move their heads and eyes about in the illuminated field between discrimination responses.

SUBJECTS

The subjects used in this experiment were seven young Rhesus monkeys (*Macaca mulatta*), two human adults, and one human child. The human subjects were Richard, a graduate student in psychology who had experienced considerable practice in brightness discrimination through use of the Macbeth Illuminometer previous to and during this experiment; Robert, also a graduate student in psychology who had had no particular practice in a brightness-discrimination experimental set-up; and Mary, a girl of seven years who was in the second grade.

Three of the seven animal subjects were used for preliminary standardization of the apparatus and were thereafter discarded. Four new animals were used in the experiment proper, two being used just beforehand for the investigation of some problems in methodology. In this section are given excerpts from the protocols of each of the four animals used in the experiment proper, and then an account of the care and handling of the monkeys.

Protocols.

Hector. Male. Weight 9 lbs. Estimated age 3-4 years. Hector was purchased from Louis Ruhe Inc., at whose animal farm he had been kept for an unknown per-

iod. There he had experienced good handling, for when brought to the laboratory he was very tame. He was introduced into the apparatus on the third day after his arrival, and became a good worker in a short time. He was active while in the reaction cage between discrimination responses, though usually attentive at reaction times. His responses were clear cut, as he withdrew into the reaction cage immediately after making his response. He appeared to be in excellent health throughout the experiment and showed no signs of loss of motivation or interest. He was more docile when handled by the experimenter than any of the other animals.

Dido. Female. Weight 10¼ lbs. Estimated age 4-5 yrs. This animal had been in the laboratory since June 1933, but had not been handled before November of that year. She was just at the beginning of adolescence as indicated by the two irregular menstrual periods which occurred between April 24 and 26 and between July 19 and 20. These were the first observed. Her performance during the menstrual periods suffered in no observable manner in the experimental apparatus. Dido was given to excessive movement, feinting behavior, and general activity which would indicate wildness. In the apparatus her responses were the least clear cut of all the animals. She attempted to touch both glass doors in quick succession, or to reverse her choice in case of an error. She tried to lean far out of the reaction-cage when making a discrimination. Early in the experiment it was suspected that this animal was responding to the illumination within the incentive-box. Modifications of the box were made to decrease the reflection therein, and to limit the animal's view to little more than the stimulus patches. Though disturbed at first by these modifications, the animal soon regained her former accuracy. Her responses became more clear cut for a time, though toward the end of the experiment they again lost their distinctiveness. It was necessary to force the shutter down on Dido's head in order to induce her to withdraw

into the reaction-cage after making a response. This procedure did not seem to hurt her. She never permitted handling by the experimenter. While she appeared to be a nervous animal, wild and untractable, she showed as low limens as any animal in the group. Her motivation was good throughout, except for a series of eight experimental periods at the beginning of the testing on Brightness Level III when she showed an aversion to work.

Menelaus. Male. Weight 7½ lbs. Estimated age 3 years. Menelaus was purchased from Louis Ruhe, Inc., being obtained soon after his arrival from the wild. When brought to the laboratory he experienced his first handling. He appeared tractable, and after two weeks preliminary training was introduced into the apparatus. Because of his relatively small size it was necessary to introduce an auxiliary platform on which he could stand in order to reach easily the shutter opening when in the standing position. He adapted speedily in the preliminary training. During the experimentation he remained quiet between discrimination responses, and responded quickly on the opening of the shutter. After 20 experimental periods on Brightness Level I he lapsed into a position habit of irregular nature, changing his preferences from day to day. Though he showed improvement after 10 experimental periods he never regained his former accuracy before the experiment was discontinued. Menelaus always appeared timid and did not allow direct handling by the experimenter.

Ulysses. Male. Weight 6½ lbs. Estimated age 2½-3 yrs. Ulysses was obtained from William Bartel, dealer, being brought to the laboratory soon after he came from the wild. He was the youngest subject and the last one brought to the laboratory. His initial adaptation required three weeks, and his preliminary training in the apparatus required more time than any of the other animals. His performance was at first slow, but became rapid and clear cut, though somewhat lacking in accuracy. He ex-

perienced more difficulty than any of the others in transferring from one brightness level to another, and since he lost so much time, was used on only levels II, IV, and III, in that order. In the apparatus he sat quiet, though apparently "suspecting the worst." Because of his small size he too was used on the auxiliary platform in the reaction-cage. He appeared in good health throughout.

Care and Handling of Animal Subjects. The animals were housed throughout the experiment in the primate section of the vivarium of the laboratory. Each was kept in a separate cage, 35½ x 27½ x 30¼ ins., made of galvanized iron and fitted with an iron-mesh door extending over the whole front. In the rear, just in front of a grilled window, was a wooden shelf for perching. The cages were equipped with moveable wire floors, with pans thereunder containing sawdust. The cages were cleaned and sawdust removed and changed twice a week by the laboratory attendant. The vivarium faced north and had no direct sunlight. The animals enjoyed, however, a half-hour exposure to a General Electric Sunlamp (Model B, vertically adjustable, type S-1 bulb), each day.

The typical daily routine for the animals is outlined below. During the summer, from July 10 to August 15, only one experimental period was used. The routine was the same as that followed during the winter except for the omission of the afternoon periods of experimentation. The schedule follows.

7:00 a.m.	Cage cleaning and exercise.
8:30-11:00 a.m.	Morning experimentation. Morning feeding immediately after each animal had finished with his experimental period.

12:00 Noon feeding.
2:30-5:00 p.m. Afternoon experimentation.
5:15 p.m., Evening feeding.

The diet of the animals was one that has been standardized in the Columbia Laboratories and has been found satisfactory for keeping the Rhesus monkey in good health over a number of years. For a more detailed account see Fjeld (3). At the morning feeding every day the animals each received a half pint of warm milk mixed with eggs in the proportion of two eggs to a quart of whole milk. At the noon feeding they received a piece of whole-wheat bread, an Irish or sweet potato, or an equal amount of boiled rice. These were given on regular days, the bread on Sunday, Monday, Wednesday and Friday, the Irish potato on Tuesday, the rice on Thursday, and the sweet potato on Saturday. Also at the noon meal the monkeys received small pieces of carrot or beet, and a lettuce leaf or piece of celery. Each animal showed an individual preference for the greenstuffs and was fed accordingly. For the evening feeding each animal was given a 3-4 oz. piece of banana with skin, sliced and sprinkled with a mixture of bone meal and Adex tablets, the latter containing cod liver oil. The condition of the faeces was used as an index of the suitability of the diet and general health of the animals. No water was supplied in the cages as this was found unnecessary. Occasionally, on very hot summer days, the animals were given small quantities of water with their evening meal.

On arrival in the laboratory each animal was fitted

with a leather belt to which a light chain was attached. For preliminary handling the leather belt was placed around the neck to insure more direct control of the animal by the experimenter. After isolation in a cage with regular feeding but without handling for 48 hours, the taming process was begun. The animal was led into the runway and into the adjoining room where it was fed raisin, apple or peanut or one of its regular meals by the experimenter. An increase in the excursion through the laboratory and a lengthening of the handling period was made on successive days. When the animal could be led without protest and could be fed manually he was considered ready for introduction into the transfer truck.

The transfer truck consisted in a closed wooden box, 11 x 13 x 10 ins., mounted on legs 45 ins. from the floor, and these on rubber tired casters. Entrance was gained through a vertically sliding door, covered with wire mesh. The box was closed otherwise. The truck was baited inside with food and the animal was induced, by coaxing or pulling the leash, into the truck. This was accomplished usually at the evening feeding, when it was required of the animals that they enter the truck in order to obtain pieces of banana. The monkeys were trained to hop directly from their living cages or from the floor into the truck. When proficient at this the animals were considered ready for introduction into the experimental apparatus.

The leather belt was transferred from the neck to the waist and the chain attached thereto was shortened to a 6 in. length, and fitted with a ring on the end. To

this ring a longer leash was snapped for handling. After the animals were well accustomed to the transfer truck and to the experimental apparatus, the belts were removed and the animals left entirely free. It was found unnecessary to handle the animals directly during the entire course of the experiment proper.

Preliminary Procedure

In this section we will describe the training of the animal subjects in the apparatus, the training in the discrimination response, and the results of some preliminary experimentation with two of the four animals used in the experiment proper.

Each animal, when first brought to the apparatus in the transfer truck and introduced into the reaction cage, found one of its regular meals therein. The room had been darkened, the dome light turned on, and the blower fan set in operation. After one to four days' feeding in this manner the animal was introduced into the empty reaction-cage from the transfer truck. The shutter opening on the forward end of the reaction-cage was opened and the experimenter gave the animal food piece-meal through the shutter opening in front of the incentive-box and the illuminated stimulus patches. The animal was trained to approach the opening as soon as the shutter began to move up and to reach for food. The food was first handed to him and later placed on the ledges in front of the glass doors of the incentive-box. After the animal became adept at reaching for food on the ledges, the doors were placed ajar and held so by a piece of food. The

monkey learned to reach for the partially invisible food and to become accustomed to the motion of the glass door as it closed when the animal's hand was withdrawn from the incentive-box. Then, by degrees, the food was placed further and further from sight toward the bottom of the incentive-box on the food tray therein. Finally it was placed directly in the tray and the glass doors closed in front of it. If the animal would not push open the door to reach for the food, as was the case in all instances, the experimenter opened it manually in the view of the animal. After 10-20 such tuitional trials the animals learned to reach through the glass doors for the invisible food. During this training both doors were unlocked and both stimulus patches illuminated. The dome light in the reaction-cage was extinguished during the time when the shutter was up and the animal was reaching for the food. After the reaching response was well established the training in brightness discrimination was begun.

At first each animal was presented with only one illuminated stimulus patch, the other one being in complete darkness. The monkeys learned to respond to the lighted side in a relatively small number of trials, ranging from 35 to 87 for the several animals. On the next step of training both the stimulus patches were illuminated, the brightness ratio of the two stimuli being 17:1. The absolute brightness of the dimmer light, the standard, was 0.7769 millilamberts, the standard brightness of level II in the experiment proper. The difference in brightness between the stimuli was

diminished on the following experimental periods. The successive steps of descent presented the following ratios of brightness between the patches: 11.8 : 1, 6.43 : 1, 5.02 : 1, 2.89 : 1, 2.68 : 1, 2.01 : 1, 1.97 : 1, 1.86 : 1, 1.73 : 1, 1.50 : 1. In some instances training was confined during a single experimental period to one brightness ratio, while on other periods two to four ratios were presented successively in descending order. When an animal had made 27/30 correct responses at any stimulus ratio that ratio was eliminated from subsequent presentation.

This training was continued with each of the animals until they were able to give 27/30 correct responses to a brightness difference in the ratio of 1.50 : 1. In the case of Hector 979 choices were made on all brightness ratios, from light-darkness down through the 1.50 : 1 ratio before this norm was attained. Dido took 857 responses to attain this criterion. Menelaus required 913, while 894 brought Ulysses to a performance of 24/30 correct choices when presented with stimuli whose brightness were in the ratio of 1.73 : 1. The modified method of limits, to be described below and which was the standard procedure for the experiment proper, was used with Ulysses at this point. He was given 786 choices by this method as further training before the experiment proper was begun with him.

In this preliminary training, as well as in the main experiment, two kinds of food were used as incentives. Small cubes of fresh apple and small raisins were

placed singly in the food tray of the incentive box. It was found in former experimentation in the Columbia Laboratory (3) that apple and raisin are among the preferred foods of the Rhesus monkey. The incentive value of these fruits remained high throughout the experiment.

The only punishment that was used for incorrect responses was the failure to get the reward. The animals were trained to withdraw immediately into the reaction-cage after touching the wrong glass door and finding it locked. Occasionally the wooden shutter was forced down on the head of an animal who did not withdraw immediately into the cage. It was found unnecessary to use any more drastic form of punishment for wrong responses.

After the training of the brightness discrimination response described above, and before the experiment proper was begun, two subjects, Hector and Dido, were used to work out certain points of experimental method. In order to determine which psychophysical method might prove most useful in this experiment, three modifications of the method of limits were tried. These are discussed in order below.

1. A series of four positions of the variable stimulus was chosen so that a brightness difference between the variable and the standard (DI/I) was represented by the following fractions: 0.60, 0.38, 0.24, 0.16. To both subjects, Hector and Dido, this stimulus series was presented in descending order eight times in a single experimental period. This is the us-

ual method of limits when approaching the limen from one side only. The results of eight experimental periods were grouped and psychophysical curves drawn relating percentage of correct choices to stimulus differences. The curves were typical of psychophysical data. The total number of correct choices made by each animal on each experimental period was about constant, indicating no disturbance to the habit during this testing. It was thus indicated that this form of the method of limits might be used. The absolute value of the limens is not useful for comparison purposes since practice had intervened between these and subsequent tests.

2. Following the tests reported above additional training was given and then Hector was tested with a second form of the method of limits. The stimulus series represented by these fractions (DI/I), 0.30, 0.23, 0.16, 0.10, 0.05, was presented to Hector, first in descending order and then in ascending order, one choice being made at each step during each series, going down four times and up four times in one experimental period. The results of ten experimental periods were used to compute the percentages of correct responses and to plot a psychophysical curve. This curve was similar to the one obtained with Hector in the test reported above. The range of stimuli was too short to give a clear-cut descent from 100 to 50 per cent correct choices. Also it was noted that the total number of correct choices in each experimental period decreased in successive periods, indicating a gradual breaking

down of the habit. It was indicated, however, that this second form of the method of limits might be applicable with animal subjects.

3. While this second method was being used with Hector, a third variation of the method of limits was used with Dido. A series of stimulus values was chosen and five choices were made by the animal at each of these steps as they were presented in descending order. All choices were made at one step before going to the next lower. This is the method used by Spence (15), and by Elder (2) on sensory limen determinations in chimpanzee. A limen whose absolute value was a good deal lower than that obtained with Dido in the first method above or with Hector in the second method which was used at the same stage of training as this one with Dido was obtained. It was indicated that this variation of the method of limits might also be used.

During this test, series of different lengths were used in order to study effect of length of series and absolute range of stimuli on the limen. The results of this test appear in Table 2 below. Limens were com-

TABLE 2

PERCENTAGE OF CORRECT CHOICES MADE BY DIDO ON A LONG AND A SHORT SERIES OF STIMULI PRESENTED IN A DESCENDING ORDER WITH FIVE CHOICES ON EACH STEP BEFORE PASSING TO THE NEXT LOWER STEP

Series	N	.63	.52	.38	.30	.23	.17	.16	.12	.10	.08	.05	.03
Short	45				84	93	84		78		77	47	48
Long	50	96	98	96	90	98		86		86		54	

The N column represents the total number of choices made on each step in each series. The results for the long series are computed from the 25 choices made on each of the long series of steps just before and just after the short series was used, in order to eliminate practice effect.

puted from this data by the quartile method to be described below (the method used in the experiment proper) by which the 75th percentile point is computed. This point was found to be .095 for the shorter series and .085 for the longer, indicating little difference in the results of the two lengths of series.

It should be noted, however, that the break in the curve for the long series is sharper and occurs on a smaller difference than that for the short series. Also, an introspection of a human subject is relevant at this point. He stated that it was desirable to have a number of large brightness differences presented at the beginning of the experimental period, so that criteria for judging might be well established for that period. Since with the short series used above only a few judgments were made at the large stimulus differences, the long series would be preferable from this point of view.

After these preliminary tests it was evident that any of the methods tried might be used in the experiment proper. It was decided to use the last modification of the method of limits which was tried with the subject Dido, employing the long series of differences, and allowing five choices on each step before descending to the next lower one. The successful use of this method by Spence (15) and by Elder (3), together with the clear-cut psychophysical curve obtained with its use in the long series with Dido, and its convenience of application were the reasons for its adoption for use in the experiment proper. It is interesting to note that Spence obtained approximately the same limens

for visual acuity in chimpanzee by this method and by a modified method of constant stimuli.

After all four animals had been trained to meet the criterion of 27 out of 30 correct choices at a stimulus brightness difference in the ratio of 1.50 : 1, or expressed by the fraction (DI/I) of 0.50, as indicated in the account of the preliminary training, and after the above-described preliminary experimentation had been completed with Hector and Dido, all four animals were ready to start on the experiment proper. The general procedure used in the experiment proper is described in the following section.

Procedure in the Experiment Proper

Testing was done at the four brightness levels used in this experiment in the following order: II, III, IV, I, except for Ulysses who was tested in the order II, IV, III. He was not used on Level I for reasons given in the Protocols above. It was thought early in the investigation that it would be possible to test at a level brighter than IV, (standard at 55.3 millilamberts). This was found to be impractical with our photometer, and so it was decided to extend the range downward instead below the brightness of Level II, (standard at 0.7769 millilamberts). Level I was then chosen at a standard brightness of 0.0864 millilamberts. Therefore, testing was done on Level I last, being carried on during the summer from July 10 to August 15 with a single experimental period each day rather than two as at the other levels.

Experimental Routine. At the beginning of each experimental period the apparatus was set in order, the food incentive preparĕd and placed on the control table, the record sheets headed, the locking magnets tested, and the proper lights turned on. The animal was brought from the living quarters into the darkened experimental room in the transfer truck. This was rolled next to the rear of the reaction cage and the doors of both pulled up. The animal jumped immediately into the reaction cage, the door was fastened behind him, and the routine testing procedure begun, starting with the largest brightness differences. A single experimental period consisted in the animal's response to five settings of the stimuli at each brightness-difference step. There were from 8 to 10 steps in each series, so that a total of 40-55 responses was made by the animal at each period. Five responses were made to the largest difference after the descending series had been completed.

For each discrimination response by the subject the following procedure was carried out.

1. Both wells of the food tray in the incentive-box were baited. (Apple was used in the morning and raisin in the afternoon.)

2. The slides on the photometer scale were set in their proper positions.

3. The magnet switch was thrown, locking one and shutting the other glass door of the incentive-box. A click accompanied the locking which served as a ready signal for the subjects.

4. The reaction-cage shutter was opened slowly, and the dome light of the reaction cage was extinguished.

5. As soon as the shutter was opened the stopwatch was started. (Stopwatch graduated in fifths of a second.)

6. The general behavior of the subject and his choice of door was observed.

7. As soon as the animal touched one of the glass doors the watch was stopped.

8. After the animal had withdrawn into the reaction cage the shutter was let down.

9. The dome light was turned on.

10. The magnet switch was opened.

11. Observations were recorded by the experimenter.

After habituation to the experimental routine the total time taken for a single discriminatory-response procedure was about 30 seconds. Thus an experimental period of fifty discrimination responses could be completed in less than thirty minutes.

Blank forms for taking records during each experimental period were mimeographed. Space was made for 60 discrimination responses, and the blank was numbered accordingly, the rows being grouped in fives. Columns headed "Stimulus," "Position," "Time," "Score," "Comparisons," and "Remarks" were drawn. The position column was filled in at the time of mimeographing, being the same on all sheets. It contained a series of R-ight and L-eft notations in irregular order. The series contained the same number of R's and L's and did not repeat either position more than three times in succession. The R's and the L's were to designate the position of the brighter of the two patches. Recording was started in a record sheet in a different place for every experimental period so that there was a systematic variation of the position se-

quence. The "Stimulus" column was filled out just before the experimental period with the value of the successive steps of brightness difference used at each step of the level being tested. Entries were made in the other columns after each response, the discrimination time in seconds going in the "Time" column; a "C" or "W" for a correct or wrong response; a number indicating the number of head and body movements made by the animal back and forth in front of the stimuli was entered in the "Comparisons" column; and incidental data were recorded in the "Remarks" column.

At the end of the day the individual records for each subject at each experimental period, morning and evening, were transferred in condensed form to a permanent record sheet for error scores. A typical record for ten experimental periods with one animal is shown in Table 3. Each row represents the number of correct responses on each stimulus step for each experimental period. It was the practice to give five choices on the largest difference after the smallest difference. This was done to check up on the stability of the habit after failures on the subliminal stimuli, and to afford additional practice. The records for these are found in the second entries in the first column for each experimental period. Records such as the one in Table 3 were kept throughout the experiment in order to study the effect of practice and to determine how many experimental periods should be devoted to a level before the performance of the subjects no longer showed im-

TABLE 3

Typical Condensed Record of the Experimental Periods for Subject Dido on Brightness Level IV, (Standard at 55.3 Millilamberts)

Period No.	Time			0.62	.51	.39	.31	.22	.15	.10	.05	.00
225	Sun.	6/17	AM	5 5	5	5	4	5	4	4	2	4
226			PM	5 5	5	5	4	4	5	4	4	4
227	Mon.	6/18	AM	5 5	5	4	5	4	5	4	4	4
228			PM	5 5	5	5	5	5	5	5	4	2
229	Tues.	6/19	AM	5 5	5	5	5	5	3	4	5	2
230			PM	5 5	5	5	5	5	5	5	4	2
231	Wed.	6/20	AM	5 5	5	5	5	5	5	5	2	2
232			PM	5 5	5	5	5	4	4	3	3	2
233	Thur.	6/21	AM	5 5	5	5	5	5	4	4	5	2
234			PM	5 5	4	5	4	4	4	4	2	4
	Totals			50	49	49	47	46	44	42	35	28
	Percentages of correct			100	98	98	94	92	88	84	70	56

In the first row at the top is given the series of brightness differences used as steps in the modified method of limits, (DI/I). On the rows below are given the number of correct choices out of five made on each step at each experimental period, including the five made at the largest difference at the end of the period (these not counted in totals).

provement due to practice. In the chapter on results, records similar to the last row of this one (marked percentages of correct), but based on 100 rather than 50 discriminations at each step, are presented. The limens are determined from the results of the last 100 responses made at each step on each level. Between 150 and 250 responses were made to each step difference at each level, during from 30 to 50 experimental

periods. Experimentation at any level was continuous with one exception.

The preliminary training and preliminary experimentation with all animals were completed about the middle of April, 1934. The experiment proper was begun immediately thereafter and continued uninterruptedly, seven days a week, with two experimental periods each day, until June 28. After nine days' interruption it was continued with experimentation once daily until August 15. As soon as experimentation was completed at one level the subjects were transferred immediately to the next. They were given sufficient preliminary training on that level to bring them to a performance of 27/30 correct choices on the greatest stimulus difference used at that level. The transfer was apparently easy for all animals except Ulysses. Hector, Dido, and Menelaus came up to the criterion on the second or third experimental period of preliminary training on the new levels, in all cases involving less than 100 discriminations. Ulysses required some 400 discriminations to move from Level II to Level IV, as this was the order in which he was tested.

Experimental Controls. In a discrimination experiment of this kind, certain systematic errors might arise. These may be of two types: first, those arising from a subject's response to any cue other than the systematic variation of the stimulus lights; and second, those arising from the subject's response to luminous intensity rather than to brightness. The former sort of error might arise in any form of discrimination ex-

periment. The latter is peculiar to certain visual discrimination situations.

Concerning the first type of error, the experimental set-up was so arranged as to enhance the value of the visual stimuli. Perhaps the best evidence that the subjects were responding to the light stimuli is the fact that their responses came to be made on a chance basis when the brightness differences approached zero.

It seems unlikely that visual cues could have been given by the experimenter. He always stood in the darkness, though in view of the subject, who had to look beyond the stimulus lights into the shadow in order to see him when making a response. The one-way light screen, and additional screens placed between the reaction-cage and the experimenter, prevented the subject from receiving any visual cue from the experimenter between responses. When discriminating the animal was observed only very occasionally to look in the direction of the experimenter before the response had been made. Further, on a number of discrimination responses the experimenter had forgotten which of the stimuli was the correct one and was surprised at the animal's choice. These controls indicate that visual cues were eliminated.

Auditory cues which might have arisen from movements of the experimenter or of the apparatus in shifting the stimuli were eliminated by the adoption of a standard routine of constant movements, making false settings when no movement was necessary, etc. It was reported by human subjects that the electric fan served

as an adequate sound screen for masking the noise of shifting stimuli.

Concerning the second possible type of error there are two controls which have to be exercised. These will be noted below. Johnson has pointed out the two different aspects of the visual stimulus which vary simultaneously in this type of visual situation. Johnson has characterized these as follows:

> In the accepted photometric nomenclature, if the light coming from a source produces on a surface the same illumination as that produced by a standard source at the same distance, both sources are said to have the same *luminous intensity* in the directions respectively given. The standard measure of luminous intensity in America is the *candle*. The *brightness* of a source is taken as the luminous intensity divided by the area of the source; it is usually measured in terms of candles per square meter or derivatives thereof. [millilamberts in this experiment]
>
> The opal glass screens limited by the windows therefore serve (1) as sources, each having a certain candle power, and (2) as visual objects, each having a certain brightness and area. (9, p. 322)

Now since in this experiment the stimulus patches of flashed opal glass have equal areas the luminous intensity and brightness of each vary simultaneously with the movement of the lamp carriages in the photometer. Since brightness vision is being studied here it must be determined that the subject is not responding to luminous intensity. In the first place, he might make his response to the luminous intensity of the stimulus patches in terms of the light which they allow to pass into the incentive box and there be reflected. This

possibility was guarded against by limiting the subjects' view through the clear glass door of the incentive-box to little more than the stimulus patches themselves, and by covering those parts which were visible with non-reflecting black velvet cloth and cardboard. (See Dido's protocols for an illustration of an animal's response to illumination of the incentive-box.)

In the second place, the subject might respond to the stimulus patches themselves as sources of luminous intensity and not as surfaces of brightness. In human experimentation the terms are often confused and in the experimental set-up the subject is allowed to make his discrimination on the basis of either of these aspects of the stimulus without subsequent determination of which cue was used. It was thought best in this experiment to vary each of these aspects independently in order to see if a subject's discrimination response was impaired. In a crucial test it was necessary to reverse the luminous-intensity and brightness cues. Such a test was performed with one of the animal and one of the human subjects.

The set-up of Brightness Level II was used for this control test, all conditions being as in the usual experimental routine except that the brightness and luminous-intensity cues were reversed. This reversal was easily accomplished by occluding one half of one of the two stimulus patches, thus reducing the amount of light emitted by one half or cutting its luminous intensity in half. Therefore, when a brightness difference expressed by a fraction (DI/I) less than 0.5 was made between the stimuli, the brighter light actually emitted

the less light, reversing the luminous intensity and brightness of the stimuli.

Hector was introduced into the control-test situation after completion of all testing done in the experiment proper. When he was required to discriminate a difference of 0.63 at the level II standard with the cues reversed he was at first disturbed and required 77 choices to attain a norm of 9/12 correct choices. A few hours later the animal was tested in a regular experimental period with a descent from greatest to least brightness difference according to the modified method of limits adopted for the regular routine. The luminous-intensity and brightness cues were reversed as indicated. A few large brightness differences were added at the start. As soon as the first few choices were made in this series the animal seemed to behave with more confidence and assurance and before the series was completed he was reacting in a manner similar to that shown on regular experimental periods. The results of this test appear in Table 4.

TABLE 4

RESULTS OF THE CONTROL TEST FOR LUMINOUS INTENSITY VERSUS BRIGHTNESS WITH HECTOR

Stimuli	*1.00*	*.80*	*.63*	*.52*	*.39*	*.30*	*.23*	*.16*	*.10*	*.05*
No. of correct choices	5	5	12* 5*	5	5	4	4	5	3	2

Showing the number of correct choices made by Hector at the end of the control experiment at Brightness Level II when the luminous-intensity and brightness cues were reversed.

*Note that there were 15 choices made on step .63 and 5 on all the rest, and on .63 after completion of the series.

It is seen that this is a typical record for any single experimental period, with all the choices correct on the

large differences and some errors on the smaller differences. This table may be compared with the rows of Table 3. Robert, one of the human subjects, was given the same test at the end of his testing on the experiment proper. While he reported some difficulty in discriminating in the early part of the test, his performance was typical of that of the regular testing periods. Warden (17, p. 76) has already indicated that some disturbance of the discrimination habit might be expected when the subject is presented with the altered situation of the control test. Though this disturbance was apparent at first with both the animal and human subject it rapidly disappeared. It is concluded from these two tests that a reversal in the luminous-intensity and brightness cues in this experimental situation had no real detrimental effect on the discrimination habit.

III

RESULTS

The results of this experiment include two types of scores—error scores and time scores. The error scores represent the percentage of correct responses which any subject made to a given difference in brightness over an arbitrary number of discrimination responses made. The time scores represent the average time in seconds taken by the subject to choose as the brighter one of the two stimulus patches of lights presented. These averages, also, are based on an arbitrary number of responses made. The error scores are the more important of the two types since it is upon the average value of these that discrimination limens are computed and comparisons made. Time scores are of interest for studying special aspects of the problem, though they furnish only supplementary data for the main question of the investigation. The two types of scores show some correlation, but cannot be used interchangeably.

In this chapter the results of the animal experimentation are given more space than those obtained with human subjects. The primary purpose of the experiment was a determination of the limens for brightness discrimination in the monkey. Limens were obtained for man only for purposes of rough comparison. Thus the results of human experimentation are considered of secondary importance and are consequently given less space. One of the main purposes in using human subjects was to determine the reliability of the photometer as an instrument for measuring brightness sensi-

tivity. This has already been suggested in the section on apparatus above (see pp. 83-84).

It will be convenient to present the results in the following order:

a. Time scores for the monkeys.
b. Error scores for the monkeys.
c. Error scores for the human subjects.
d. Treatment of error scores and discussion.

TIME SCORES FOR THE MONKEYS

The time scores for the monkeys represent the average discrimination times, taken from the time the shutter of the reaction-cage went up until the animal had touched one of the glass doors, for each animal at each stimulus step on the brightness levels considered. The scores for Hector, Dido, and Menelaus are grouped together for analysis. A separate frequency distribution was made for the correct and for the wrong responses at each step, and the averages and their probable errors computed for each. These are presented in sections *A* and *B* of Table 5. Also the differences between discrimination times for the correct and wrong responses for each step are given together with the reliabilities of these differences. The average discrimination times for the correct responses for both levels are plotted against the fraction representing brightness difference at each step for each level in the two parts of Figure 4. From the table and the figure we are able to study three aspects of the results.

First, we observe that the successive average time

A. Brightness Level I.

Step	Correct No.	Correct Av.	Correct $PE_{av.}$	Wrong No.	Wrong Av.	Wrong $PE_{av.}$	D.	$PE_{diff.}$	$D/PE_{diff.}$	Chances in 100
1.55	442	2.84	.0408	35	3.67	.310	.83	.313	2.65	96
1.25	454	2.76	.0601	45	3.41	.100	.65	.119	5.4	100
.90	515	2.97	.0428	64	2.55	.143	—.42	.149	2.8	97
.65	480	3.04	.0527	85	3.95	.185	.91	.192	4.73	100
.51	436	3.56	.0707	91	3.50	.143	—.06	.159	.38	59
.44	453	3.59	.0735	117	3.76	.165	.17	.180	.94	74
.37	427	3.59	.0730	141	4.06	.147	.47	.164	2.86	97
.30	384	3.65	.0704	168	4.44	.157	.79	.174	4.5	100
.23	366	3.81	.1026	184	3.98	.145	.17	.177	.99	74
1.55 (end)	334	2.87	.0384	23	2.50	.523	—.37	.166	2.23	93
Averages		3.27			3.58	.202	.31	.179	1.72	87

B. Brightness Level II.

Step	Correct No.	Correct Av.	Correct $PE_{av.}$	Wrong No.	Wrong Av.	Wrong $PE_{av.}$	D.	$PE_{diff.}$	$D/PE_{diff.}$	Chances in 100
.63	530	2.55	.0480	26	3.23	.205	.68	.210	3.2	98.5
.52	540	2.54	.0447	23	3.93	.399	1.39	.404	3.44	98.9
.38	529	3.15	.0490	36	4.00	.218	.85	.214	3.98	99.7
.30	497	3.18	.0417	58	4.32	.218	1.14	.218	5.23	100
.23	459	3.41	.0605	85	4.94	.255	1.53	.262	5.8	100
.16	372	3.55	.0598	110	4.42	.150	1.10	.161	6.9	100
.10	374	3.57	.0544	147	4.65	.177	1.08	.185	5.8	100
.05	290	3.79	.1160	180	4.96	.154	1.17	.186	6.3	100
.63 (end)	478	3.26	.0450	21	2.40	.232	— .94	.236	3.98	99.9
Averages		3.20	.		4.10	.225	.90	.231	3.91	99.6

The steps are marked in fractions of brightness difference. The No. columns give the number of discrimination responses correct or wrong at each step. The differences are marked positive when the wrong is greater than the correct.

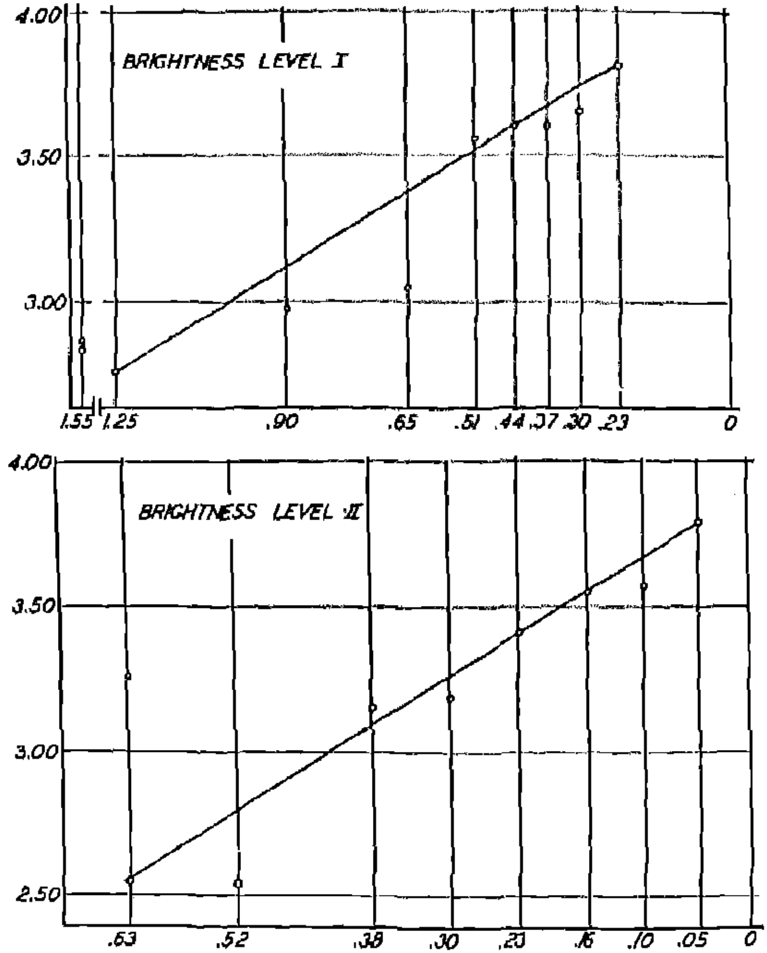


FIGURE 4

AVERAGE DISCRIMINATION TIMES ON CORRECT RESPONSES FOR THREE ANIMALS (HECTOR, DIDO AND MENELAUS), ON BRIGHTNESS LEVELS I AND II AS RELATED TO STIMULUS DIFFERENCE

The ordinate is laid off in seconds and the abscissa in fractions of brightness difference. The circles mark the averages for the descending series, and the triangles the average for the final large brightness difference.

scores for the correct discriminations at both levels, and for the wrong ones at level II, seem to increase in magnitude as the stimulus difference is decreased. The difference in average times between any two adjacent steps is not significant except in a few instances. The differences between the extremes, however, are reliable. Also, the trend of the increase in time is more or less *regular as indicated by the empirical straight line* which is fitted to the points plotted in Figure 4. The discrimination times for the large stimulus differences presented at the end of each experimental period are in three instances of about the same magnitude as the times for the difference when presented initially. These times are represented by triangles in Figure 4. The average time for step .63 (end), correct, at level II is intermediate between the smallest and the largest time scores at that level. The relatively small time scores for these large differences presented at the end indicate that the increase in discrimination time is a function of the stimulus difference, and not of some other factor resulting simply from the serial order of presentation of the stimuli.

Second, we observe that the average discrimination *times for the correct responses at both levels* are smaller than those for the wrong responses. The difference at level I is 0.31 sec., with a probable error of the difference of .18, giving 87 chances out of 100 that this is a true difference. For level II the difference is .90 with a probable error of the difference of .23 giving 99.6 chances out of 100 that the difference is a true one. Examination of the differences between correct and

wrong times at level I indicates that there is no consistent trend, there being three cases in which the wrong response actually took less time. It is seen that the differences are not entirely reliable on the seven steps at level I and on four steps at level II. The differences for each step at level II are all in favor of the longer times for the wrong responses except in the 0.63 (end) step, where the difference is reversed. It is interesting to note that the wrong responses on the final step of both levels took less time than did the correct ones. This may indicate that more "snap judgments" were made at these steps than at the others. The lack of any consistent trend in the differences between times for correct and wrong responses along the stimulus series prevents us from drawing any inferences as to the type of judgment made at the other steps. Probably the relative amount of comparison behavior shown by the monkeys at the various steps may account for a large part of the differences in discrimination times. In general, it was found that the number of alternations or comparison movements made by the animal back and forth in front of the stimuli correlated roughly with the discrimination time. Distractions, of course, also tended to increase the discrimination times. Third, we observe that the average discrimination times for the correct responses on both levels I and II are very nearly equal, 3.20 and 3.27 sec. Since level II was tested first and level I last, this indicates that there was no apparent change in discrimination times for correct responses during the experiment proper. In the case of the wrong responses the average time for level II is

actually shorter than for level I by 0.52 sec. The probable error of the difference here, however, is 0.30 so that the chances are only 88 out of 100 that this difference is a reliable one. The results show, then, that these animals must have been working at about constant speed throughout the experiment proper. This also suggests that the motivation was relatively constant.

ERROR SCORES FOR THE MONKEYS

The error scores for the monkeys are in terms of the *percentage of correct responses made at each stimulus* difference by each animal. These percentages are calculated from the data on an arbitrary number of discrimination responses at each step, either 50 or 100, and are presented for each animal at each brightness level in Tables 6-9. Empirical psychophysical curves were drawn for each subject representing the percentage of correct responses made at each stimulus difference of any brightness level on the last hundred responses made at that level. These are represented in Figures 5-8. It will be convenient to discuss the tables and figures for each level in separate sections below. The results will be examined for evidences of practice effects, irregularities, amounts of variability, and special circumstances in the experimental procedure of each level will be noted.

Brightness Level I. The data for level I are found in Table 6 and Figure 5. The table may be examined for practice effect. Hector showed some improvement between the first and second group of responses. Dido's improvement was much more apparent; so much so, in

TABLE 6

PERCENTAGE OF CORRECT DISCRIMINATIONS MADE BY EACH ANIMAL TESTED ON BRIGHTNESS LEVEL I, WHOSE STANDARD WAS 0.0864 MILLILAMBERTS

Subject	Series	1.55	1.25	.90	.65	.51	.44	.37	.30	.23	.16	.08	.00
Hector	1	91	91	88	87	79	74	70	72	60	51*		
	2	93	93	92	89	82	79	74	61	65	48*	53*	
Dido	1	93	92	95	87	91	76	83	79	67			
	2	94*	100*	96	98	95	92	86	77	76	69	59	
	3	94*	98*	98*	96*	94*	94*	88*	76*	80*	76*	56*	62*
	L.H.	96	99	97	97	94	93	88	79	80	73	55	62*
Menelaus	1	92	86	79	80	76	75	68	70	55			
	2	92*	96*	84*	84*	76*	70*	68*	72*	54*			
	L.H.	90	89	82	83	74	73	63	71	54			

The above data are grouped for convenience into series of 100 discrimination responses each, except in those cases marked by an asterisk (*) in which only 50 discriminations were given. The first row represents the brightness difference at each step divided by the standard of this level (DI/I). On the rows below are given the percentage of correct responses for each animal for each series of 100 or 50 discrimination responses. In those instances where the last numbered series is based on only 50 discrimination responses a row marked L.H. is added, giving percentages based on the last 100 responses made, since the last hundred is to be used in the calculation of limens. A series of 100 discriminations represents 20 experimental periods.

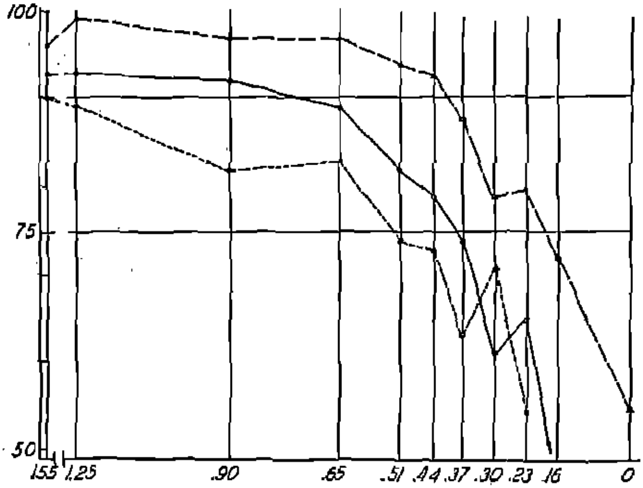


FIGURE 5

PERCENTAGES OF CORRECT RESPONSES MADE BY EACH ANIMAL SUBJECT TESTED ON BRIGHTNESS LEVEL I (STANDARD AT 0.0864 MILLILAMBERTS)

The abscissa is laid off in fractions representing the brightness difference at each step divided by the standard of this level, (DI/I). The ordinate is laid off in percentage of correct responses, from 50 to 100. The line for each animal is represented as follows:

Hector	solid line
Dido	broken line
Menelaus	dotted line

The percentages used are those for the last 100 discrimination responses made by each animal on this level as indicated in Table 5.

fact, that she was given an extra group of ten experimental periods, *50 discriminations on each step*, in order to eliminate this practice effect. Menelaus, as noted in the Protocols in the previous chapter, developed a series of position habits after the twentieth experimental period. The effects of this disturbance were never fully eliminated, and while the percentages

shown for the second series are about as high as those for the first, there is no evidence of improvement.

Examination of the table and the curves drawn therefrom in Figure 5 indicates a marked variability among the three subjects tested. It will be seen that the variability here is greater than at any other level. The reasons for this are difficult to determine. Possibly the motivation was different for each subject. Some disturbing influence must have affected Menelaus to have produced the position habits. This influence may have been in operation during the entire testing at this level. The sharper turn of the psychophysical curve for Dido, as compared with the other subjects,

TABLE 7

PERCENTAGE OF CORRECT DISCRIMINATIONS MADE BY EACH ANIMAL TESTED ON BRIGHTNESS LEVEL II. STANDARD AT 0.7769 MILLILAMBERTS

Subject	Series	0.65	.52	.38	.30	.23	.16	.10	.05	.03	.02	.00
Hector	1	96	94	94	86	84	81	64	57			
	2	94	92	93	87	80	80	82	73	57		
Dido	1	94	99	96	90	87	85	77	53			
	2	96*	98*	96*	90*	98*	86*	86*	54*			
	L.H.	95	99	96	93	95	98	82	55			
Menelaus	1	95	98	90	89	77	64	63	51			
	2	91	98	96	90	87	75	71	67			
	3	98*	98*	92*	94*	88*	80*	72*	66*	64*		34*
	L.H.	96	99	94	95	87	81	74	63	64*		34*
Ulysses	1	94	93	87	87	75	69	66	64			
	2	90	96	97	92	85	82	70	49			

The above data are grouped for convenience into series of 100 discrimination responses each, except in those cases marked by an asterisk (*) in which only 50 discriminations were made. The first row represents the brightness difference at each step divided by the standard of this level, (DI/I). On the rows below are given the percentage of correct responses for each animal for each series of 100 or 50 discrimination responses. In those instances where the last numbered series is based on only 50 discrimination responses, a row marked L. H. is added, giving the percentages based on the last 100 responses made, since the last hundred is used in calculation of limens. A series of 100 responses represents 20 experimental periods.

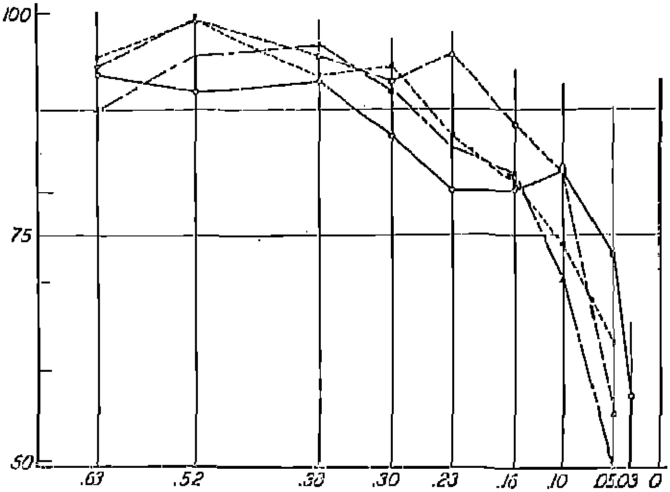


FIGURE 6

PERCENTAGES OF CORRECT RESPONSES MADE BY EACH ANIMAL TESTED ON BRIGHTNESS LEVEL II (STANDARD AT 0.7769 MILLILAMBERTS)

The abcissa is laid off in fractions representing the brightness differences at each step divided by the standard of this level, (DI/I). The ordinate is laid off in percentage of correct responses, from 50 to 100. The percentages used are those for the last 100 discrimination responses made by each animal as indicated in Table 6. The graph for each animal is represented as follows:

Hector	solid line
Dido	broken line
Menelaus	dotted line
Ulysses	dot and dash line

may indicate that her general motivation was better, since she did not make nearly as many errors on the large differences as did the others.

As has already been noted this level was the last one tested. Only three animals were used, since, as noted in the Protocols, Ulysses lost so much time in trans-

ferring from one level to another that he did not get to the last level before the experimentation was discontinued. Also this level was tested during the summer months with only one experimental period each day. Since the experimental conditions of this level are different from those at the other three levels, the difference in variability and absolute performance between the results of this and other levels may be in part a function of the slightly different conditions.

Brightness Level II. The data for this level are presented in Table 7 and Figure 6. It is seen that the range of stimuli here is smaller than that of the preceding level, and is about the same as those of the two following levels III and IV. As noted in the section on calibration, the range of stimuli chosen for each level was the result of preliminary experimentation to determine roughly the magnitude of the limen at that level.

The indications of improvement during the testing procedure at this level are most apparent on the stimulus steps presenting the smallest brightness differences. These stimuli at the end of the series of decreasing brightness differences will be called hereafter the "critical stimuli," since it is the performance on these stimuli which most influences the magnitude of the limens. It will be observed that the range of the stimulus steps showing the most improvement is greatest in the case of Menelaus and of Ulysses.

It is apparent from the table, and more especially from the graph, that the variability between subjects at this level is a great deal less than that on level I. All

the psychophysical curves cluster quite closely together and cross each other at various points. The curves show a rather gradual bend in the middle rather than a sharp break. The only important inversion is that shown in Hector's curve. It is interesting to note that all the subjects except Hector made a higher percentage of correct choices on the second than on the first stimulus step. This apparently indicates that there was some improvement due to initial adjustment at the beginning of each experimental period. These initial upturns in the curves indicate that perhaps it would have been wise to have given a few stimuli at the be-

TABLE 8

PERCENTAGE OF CORRECT DISCRIMINATIONS MADE BY EACH ANIMAL TESTED ON BRIGHTNESS LEVEL III, STANDARD AT 7.146 MILLILAMBERTS

Subject	Series	0.65	.53	.44	.36	.30	.24	.18	.14	.10	.05	.00
Hector	1	94	98	95	94	91	88	79	78	64	55	
	2	92*		88*		86*	92*	86*	74*	80*	50*	
	L.H.	92		91		87	90	84	77	71	54	
Dido	1	96	96*	94	94*	91	90	82	75	64	63	
	2	98		96		91	90	85	83	74	67	60*
	3	100*		96*		92*	92*	90*	80*	64*	46*	56*
	L.H.	98		97		93	91	88	83	69	58	58
Menelaus	1	96		95		87	87	90	77	61	67	
	2	96		95		94	93	84	80	64	37	50*
Ulysses	1	92		96		85	85	73	78	71	61	40*
	2	80*		98*		94*	84*	86*	78*	78*	56*	50*
	L.H.	87		98		87	84	82	80	75	62	50

The above data are grouped for convenience into series of 100 discrimination responses each, except in those cases marked by an asterisk (*) in which only 50 discriminations were made. The first row represents the brightness difference at each step divided by the standard of this level, (DI/I). On the rows below are given the percentage of correct responses for each animal for each series of 100 or 50 discrimination responses. In those instances where the last numbered series is based on only 50 discrimination responses a row marked L.H. is added, giving the percentages based on the last 100 responses made, since the last hundred is used in calculation of limens. A series of 100 responses represents 20 experimental periods. After some experimentation it was decided to eliminate steps .53 and .36.

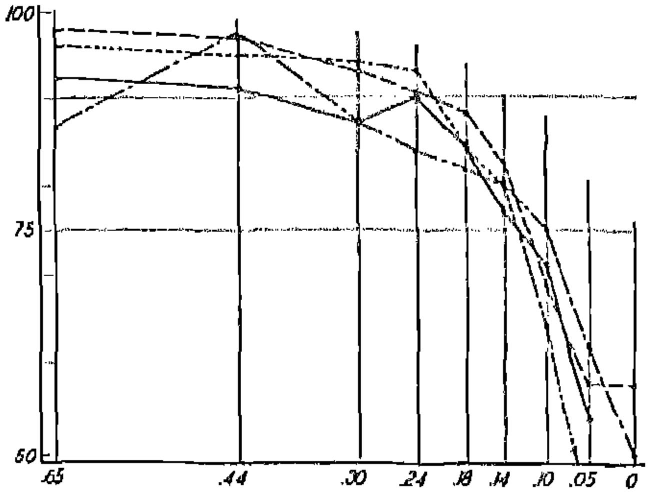


FIGURE 7

PERCENTAGES OF CORRECT RESPONSES MADE BY EACH ANIMAL SUBJECT TESTED ON BRIGHTNESS LEVEL III (STANDARD AT 7.146 MILLILAMBERTS)

The abcissa is laid off in fractions representing the brightness difference at each step divided by the standard of this level, (DI/I). The ordinate is laid off in percentage of correct responses, from 50 to 100. The percentages used are those for the last 100 discrimination responses made by each animal as indicated in Table 6. The graph for each animal is represented as follows:

Hector	solid line
Dido	broken line
Menelaus	dotted line
Ulysses	dot and dash line

ginning of each experimental period, and not included the results in the regular records, in order to allow the animals a chance for an initial adjustment each day. The same sort of upturn at the beginning of the curves will be noted at the other brightness levels.

Brightness Level III. The data for this level are presented in Table 8 and Figure 7. It will be observed

that no percentages are given for the second and fourth stimulus steps except for the first series of Hector and Dido. It was found that the experimental period was made tediously long with the inclusion of these steps. They were omitted and it was found that there was no detriment apparent in the records of Hector and Dido on the second series. Menelaus and Ulysses were tested at this level slightly later than the other subjects. The second and fourth stimuli, therefore, were never presented to them.

Practice effect is less apparent at this level than at the preceding ones. Where the percentages do show a decrease it is usually on the critical stimuli. Dido showed a temporary loss of motivation for eight experimental periods after group No. 1 of this level. She was given an extra series of ten experimental periods to eliminate any loss she may have sustained during this poorly motivated period.

The variability at this level appears to be even less than at the other levels already discussed. The psychophysical curves run very close together and cross each other at many points. The initial adaptation effect at this level is very apparent in the case of Ulysses. His percentage of correct responses is observed to increase from 87 to 98 on the first two steps. Judging from the rest of the curve, however, this second percentage may be spuriously high. The curves approach a chance value of 50-per-cent-correct discriminations at this level more nearly than do those at the second level. The inversions, with the exception of the one just noted above for Ulysses, are not large.

TABLE 9

PERCENTAGE OF CORRECT DISCRIMINATIONS MADE BY EACH ANIMAL TESTED ON BRIGHTNESS LEVEL IV. STANDARD AT 55.3 MILLILAMBERTS

Subject	Series	0.62	.51	.39	.31	.22	.15	.10	.05	.00
Hector	1	95	90	87	89	83	85	63	66	46*
	2	94	97	96	90	88	82	73	57	54
Dido	1	97	98	97	97	82	75	65	68	50*
	2	100	98	99	97	90	89	85	66	57
Menelaus	1	96	95	92	85	88	74	78	71	60*
	2	97	94	94	91	89	87	72	66	56
Ulysses	1	90	87	82	83	79	71	69	61	50*
	2	89	96	94	90	89	78	71	57	48

The above data are grouped for convenience into series of 100 discrimination responses each, except in those cases marked by an asterisk (*) in which only 50 discriminations were made. The first row above the top line represents the brightness difference at each step divided by the standard of this level (DI/I). On the rows below are given the percentage of correct responses for each animal for each series of 100 or 50 discrimination responses. The calculation of limens is done with the last 100 responses made by each animal. A series of 100 responses represents 20 experimental periods.

Brightness Level IV. Table 9 and Figure 8 present the data for this level. The practice effect at this level is slightly greater than at the other three levels tested. It had disappeared in all cases, however, before the last (fourth) group of experimental periods was given. Therefore, experimentation at this level was confined to exactly forty experimental periods, or 200 discrimination responses at each stimulus difference. The practice effect as indicated in Table 9 is most apparent on the critical stimuli. Menelaus shows the least improvement of the four subjects tested here.

The variability as indicated by Figure 8 is slightly greater at this level than at level III. Perhaps the most distinctive feature of the group of curves for this level is the rather clear break which they show at brightness differences between 0.22 and 0.10. The de-

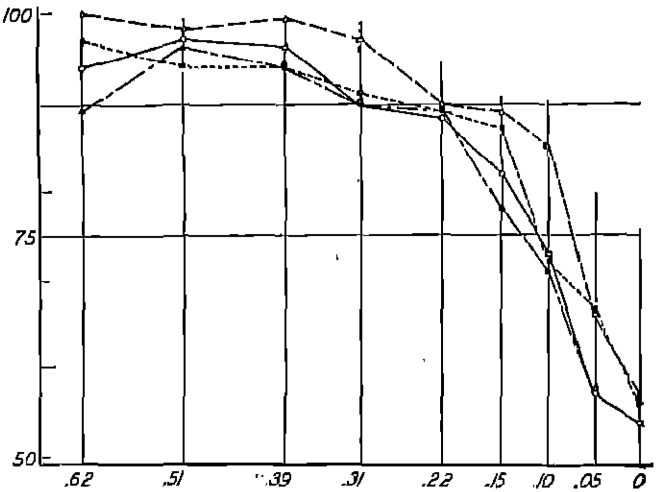


FIGURE 8

PERCENTAGES OF CORRECT RESPONSES MADE BY EACH ANIMAL SUBJECT TESTED ON BRIGHTNESS LEVEL IV (STANDARD AT 55.3 MILLILAMBERTS)

The abcissa is laid off in fractions representing the brightness difference at each step divided by the standard of this level, (DI/I). The ordinate is laid off in percentage of correct responses, from 50 to 100. The percentages used are those for the last 100 discrimination responses made by each animal on this level as indicated in Table 8. The graph for each animal is represented as follows:

Hector	solid line
Dido	broken line
Menelaus	dotted line.
Ulysses	dot and dash line

scent toward a chance performance therefrom is steep in all cases. Initial adaptation is again suggested in the case of Hector and Ulysses. These are the only inversions of much importance in the curves.

Experimentation was continued at this level down to stimuli presenting no brightness difference, marked 00, and then back to the initial difference. This had been

done at other levels only where the subjects' performance seemed to justify the continuance of experimentation at small differences and equality. It was done here with all animals as a systematic check on the discrimination response. It is apparent that without a light cue the performance of all subjects was practically chance, or 50-per-cent-correct choices. It should be noted, however, that the average of the percentage of correct choices on the 00 step for all animals is slightly above 50, being 52.25. This may indicate that there was some sort of a constant error present, affecting some of the animals more than others. This would, of course, if shown to be present, make the percentages at the other steps spuriously high. The possible source of this small constant error has not been identified. Since we are not sure of its presence, it seems best not to correct the other percentages in terms of the deviation of this average from 50 at the 00 step. Again, we may not have a large enough sample to make this deviation from chance at all significant. It is not statistically reliable.

Error Scores for the Human Subjects

Three human subjects were used for the experiment proper. Of these only one, Robert, was used at all four levels. Richard was used on the first three, and Mary on the second only. Their error scores in terms of percentages of correct responses, as well as the number of responses made at each stimulus step, are presented for all brightness levels in Table 10. Typical

TABLE 10

PERCENTAGES OF CORRECT RESPONSES MADE BY EACH HUMAN SUBJECT ON ALL FOUR BRIGHTNESS LEVELS. PERCENTAGES ARE BASED ON ALL DISCRIMINATORY RESPONSES MADE BY THE SUBJECTS

A. Brightness Level I., Standard at 0.0864 millilamberts

Subject		1.55	1.25	0.90	.65	.51	.44	.37	.30	.23	.16	.10	.05	.03	.00
Richard	N	3	3	6	6	31	28	26	33	41	40	40	35	35	40
	%	100	100	100	100	100	100	100	100	95	92	90	60	52	47
Robert	N	8	6	6	6	6	41	42	50	40	45	40	25	20	25
	%	100	100	100	100	100	100	95	94	87	82	65	52	45	52

B. Brightness Level II. Standard at 0.7769 millilamberts

Subject		.63	.52	.38	.30	.23	.16	.10	.05	.03	.02	.01	.00
Richard	N	38	35	35	37	39	71	53	60	55	50	20	5
	%	100	100	100	100	98	100	96	80	68	60	35	33
Robert	N	6	9	9	17	12	36	25	25	25	20		
	%	100	100	100	100	100	97	88	56	72	65		
Mary	N		27	27	29	31	45	45	45	40			
	%		100	100	100	93	88	68	60	57			

C. Brightness Level III. Standard at 7.146 millilamberts.

Subject		.65	.44	.30	.24	.18	.14	.10	.05	.03	.00
Richard	N	15	15	17	22	20	25	25	25	30	20
	%	100	100	100	100	90	96	76	88	66	53
Robert	N	3	19	14	21	23	25	25	30		6
		100	100	96	95	91	80	80	70		60

D. Brightness Level IV. Standard at 55.3 millilamberts.

Subject		.62	.51	.39	.31	.22	.15	.10	.05	.03	.00
Robert	N	15	15	15	19	21	25	25	25		20
	%	100	100	100	94	90	80	54	58		45

Above the top line in each section are given the fractions representing the brightness difference at each step (DI/I). In the row marked "N" for each subject is given the number of responses made at each stimulus difference. Percentages of correct responses made are based on this number of responses.

psychophysical curves, those for Brightness Level II, are given in Figure 9.

It is seen from Table 10 that a variable number of presentations of stimuli were made at the different steps of each level. This was done in order to conserve time and to obtain as many responses to the criti-

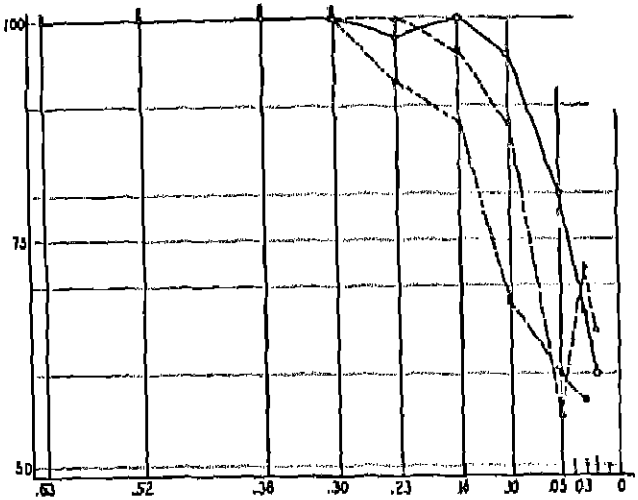


FIGURE 9

PERCENTAGE OF CORRECT CHOICES MADE BY THREE HUMAN SUBJECTS ON BRIGHTNESS LEVEL II (STANDARD AT 0.7769 MILLILAMBERTS)

The abcissa is laid off in fractions representing the brightness difference at each step divided by the standard of this level, (DI/I). The ordinate is laid off in percentages of correct responses, from 50 to 100. These percentages are based on the total number of responses given by each subject, as shown in Table 9. The graph for each subject is represented as follows:

Richard	solid line
Robert	broken line
Mary	dotted line.

cal stimuli as possible. The data are too limited to afford any opportunity to study practice effect. Had the human subjects been tested for as many experimental periods as were the monkeys they would doubtless have shown improvement. Individual differences are apparent from the table and the graph. Richard's performance on all tested levels appears to be the best

Robert's next, and Mary's last on the one level at which she was tested.

Perhaps the most obvious difference between the performance of human and animal subjects is shown by their characteristic psychophysical curves. Examination of the table and graph will show that the humans were able to choose correctly 100 per cent of the choices down to the critical stimuli. After passing a single critical point their accuracy drops sharply toward chance. The animals, as has been shown (compare Figure 6 and Figure 9), made errors even on the largest differences, and did not always show a very sharp break in their curves at the critical stimuli. The significance of this apparent difference between animal and human performance will be discussed in the following section.

As stated in the previous chapter the experimental conditions under which human subjects were tested were equated as nearly as possible with those for the animal subjects. Testing was in most cases regular, with one or two experimental periods each day. Since fewer responses were made by the human subjects their results are obviously less reliable than the animal ones. Again, practice effect was not nearly as well controlled as in the animal investigation. As noted in the section on Subjects, they had differing amounts of practice on experimental brightness discrimination. Since Richard had used the Macbeth Illuminometer for many months before this experiment began we might say that he had at least as much specific practice on a brightness discrimination apparatus as did the animals.

Robert and Mary had had definitely less. Concerning the general practice obtained in everyday life by the different animal and human subjects we can make no definite statement. The child which we used, Mary, was probably slightly younger, proportionally, than the older two animals, and perhaps fairly well matched in maturity status with the two younger subjects. Obviously they were not matched in amount of practice in everyday brightness discrimination. Therefore, any comparison of results between the two types of subjects may be taken only tentatively and roughly.

Treatment of Error Scores and Discussion

In order to render the above results for error scores more suitable for comparison with each other it has been convenient to determine a liminal point from each psychophysical curve, or from the numerical record of each subject's performance at each brightness level. Some consideration must be given to the method used in computing these limens.

It is apparent that the typical psychophysical curves which were obtained with the animal and human subjects represent half of a psychophysical curve which would have been obtained had the variable brightnesses approached the standard from below and above, rather than only from above. The curves as presented represent the percentage of judgments of the variable brighter than the standard. The percentages are relatively high at large differences and fall off rapidly at small differences, the curves sloping most steeply as they approach the equality point. These curves would

continue on if the variable had been made less than the standard and would have eventually flattened out near the zero abscissa when practically no judgments of the variable as brighter than the standard would have been made. Thus we may consider that our curves represent roughly half the typical psychophysical curve obtained in the method of Constant Stimuli in human experimentation when only two categories of judgment, "brighter" and "dimmer," are used. The procedure recommended by Jastrow (8) of choosing the 75th percentile point as the limen when equal judgments are excluded has been adopted in this experiment, since it is assumed that at this point 50 per cent of the judgments are given by chance and 25 per cent by real judgment. A convenient method of calculating these limens is the computation of the upper quartile of the distribution of percentages of correct choices, of which we have the upper half. Free-hand smoothing of the curves is used where small inversions affect the value of the 75th percentile point. Since our stimuli are laid off in fractions representing the relative brightness differences of the steps, the numerical value of the quartile, the Q_3 (75th percentile point) minus the Q_2 (50th percentile point) (since it is assumed that these *curves represent one-half of a symmetrical distribution* about the equality point), is exactly equal to the numerical value of the 75th percentile point. Further, the probable error of the quartile may be used as a measure of the reliability of the limen.

This method of calculation possesses, beside simplicity, the advantage of not weighting heavily the per-

centages of "brighter than" judgments made on stimuli showing the largest differences in brightness. It has been noted above that the animal and human psychophysical curves show a characteristic difference. This difference consists in the fact that the human subjects were able to return 100-per-cent-correct judgments on stimuli whose differences in brightness were well above those on which they failed consistently. The animals, on the other hand, rarely made perfect records even on the largest brightness differences. This characteristic difference has been shown by Spence (15). It is assumed that these early errors made by the animals are partly, at least, the result of factors of inattention, chance distracting stimuli, variations in motivational conditions from day to day, differences in the kind of instructions that could be given to the two classes of subjects, and general lack of inhibitory control on the part of the animals. Thus these errors are considered to be largely experimental errors, due more to the inability of the investigator to set the problem for the animal than to any real inability of the animal to discriminate the large differences presented early in the series. If this assumption is correct, then a method of computing limens should be chosen which does not give much weight to the early errors. The use of the quartile method suggested above meets this difficulty. Were the judgments made on both sides of the standard and the whole, instead of half the psychophysical curve obtained, some method of computation such as the Müller-Urban process might have been used (if the data were sufficiently near normal in distribution),

since in this method extreme frequencies are given much less weight than those around the 50-per-cent point.

In Table 11 are presented the limens calculated by the quartile method for each animal and human subject at each brightness level, together with the probable errors of these limens. Average limens for animal

TABLE 11

SEVENTY-FIFTH PER CENT LIMENS AND PROBABLE ERRORS OF LIMENS FOR EACH SUBJECT AT EACH BRIGHTNESS LEVEL. ALSO SHOWING AVERAGE LIMENS AND PROBABLE ERRORS FOR ANIMAL AND FOR HUMAN SUBJECTS AND DIFFERENCES AND RELIABILITIES OF DIFFERERENCES BETWEEN ANIMAL AND HUMAN SUBJECTS

	B.L.I		B.L.II		B.L.III		B.L.IV	
	L_{75}	$P.E._L$	L_{75}	$P.E._L$	L_{75}	$P.E._L$	L_{75}	$P.E._L$
Animal								
Hector	.384	.030	.061	.0048	.126	.0100	.111	.008
Dido	.180	.014	.087	.0068	.117	.0092	.073	.006
Menelaus	.525	.040	.117	.0092	.127	.0100	.110	.008
Ulysses			.125	.0098	.100	.0079	.129	.010
Averages	.363	.026	.098	.008	.117	.009	.106	.008
Human								
Richard	.075	.009	.042	.005	.069	.011		
Robert	.136	.017	.081	.012	.075	.011	.144	.022
Mary			.121	.013				
Averages (Rich. Robt.)	.105	.013	.061	.008	.070	.011		
$Av._{an'l.}$-$Av._{hu.}$	.258		.037		.047			
$P. E._{diff.}$		.029		.011		.0142		
Diff./$P.E._{diff.}$	8.8		3.3		3.4			
Chances in 100	100		99		99			

The above limens are calculated by the quartile method described in the text. Animal limens are based on results of the last hundred discriminations made at each step of each level. Human limens are based on all the data for human subjects, an average of 30 discriminations on each step of each level.

and human subjects, differences between these, and the reliabilities of the differences are also presented. If the average limens for the animal subjects as given in the table are compared with the ones which would be obtained from inspection of the average psychophysical curves given for the animal subjects in Figure 10, it will be seen that they are not greatly different.

The limens presented in Table 11 may be examined

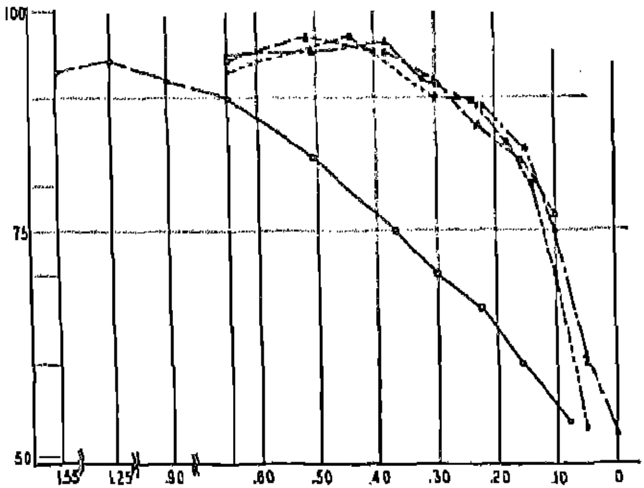


FIGURE 10

COMPOSITE PSYCHOPHYSICAL CURVES FOR ALL THE MONKEYS TESTED ON EACH BRIGHTNESS LEVEL

These curves are drawn from the averages of the percentages of correct choices made by each animal tested at each brightness level. The ordinate represents the average percentage of correct responses and the abscissa represents fractions of brightness difference. The curves for the four brightness levels are represented as follows.

Brightness Level I	solid line
Brightness Level II	broken line
Brightness Level III	dotted line
Brightness Level IV	dot and dash line

for individual differences. If the average limens for the four animal subjects are computed for Brightness Levels II, III, and IV, the following results are obtained:

Hector	.099±.007
Dido	.092±.007
Menelaus	.118±.007
Ulysses	.118±.013

It is seen that the differences between these are statistically unreliable. On Brightness Level I, however, the differences are much greater and are reliable. Dido's is the lowest, .180±.014, Hector's next, .384±.030, and Menelaus' highest, .525±.040. These differences in limens on all the levels correspond closely to the relative variabilities in the psychophysical curves considered for each brightness level above. Among the human subjects the lowest limens are found in the case of Richard, .062±.008, with Robert next, .097±.013, and Mary last on the level at which she was used. It has been noted that Richard had had considerable practice at brightness discrimination which probably accounts for his low limens in comparison with the others. The difference between Mary's and Richard's limens on level II is significant, while that between Mary's and Robert's is not, being 2.3 times the $P.E._{diff.}$ or 94 chances in 100 of being a true difference.

Turning now to the average limens for each group of subjects on the different brightness levels we find that the animals' average limen for level I is reliably higher than those for the other three levels. The

differences between the other three are not reliable. With the human subjects, the average of Robert's and Richard's limens for level I is greater than their averages for levels II and III. The differences here are reliable (99 chances out of 100). In the case of Robert the difference between his limens for level I and level IV is not reliable.

At the bottom of the table are given the differences between the average limens for the animal and human subjects (Richard and Robert) on the levels at which both animal and human subjects were tested. The average human limens are in all cases lower and the differences are reliable. If we compare the limens of Richard alone with the average of the animals the differences are even greater in favor of the human. Comparing them with Robert's, however, the differences are not so consistent nor so great. Further, if we compare Robert, the poorer adult human subject, with Dido, the best animal, we find that the human limen is lower in two instances, almost equal in one, and greater in one, indicating very little difference over the whole range. Mary's limen on level II is higher, though not reliably so, than the average for the animals at that level. From these comparisons, then, it is apparent that the animal limens seem to be definitely higher on level I than the human ones. On the other levels the difference is still slightly in favor of the humans but not nearly as great. The average differences are in all cases reliable.

The average limens for the animal and human subjects have been plotted against the standard brightness

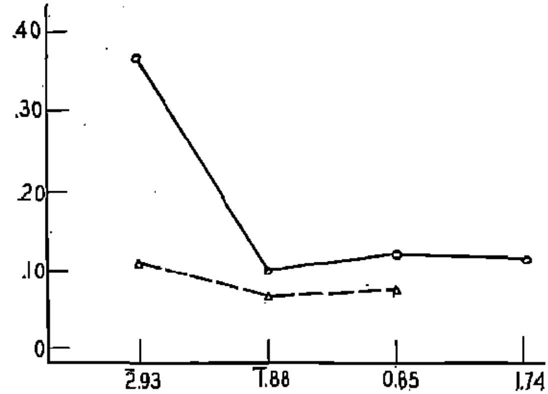


FIGURE 11

SHOWING THE LIMENS FOR THE FOUR BRIGHTNESS LEVELS FOR ANIMAL AND HUMAN SUBJECTS AS GIVEN IN TABLE 11. The ordinate represents fractions of brightness difference, (DI/I), and the points on the abcissa represent the standard brightnesses of the four levels expressed as logarithms of their values in millilamberts. The solid line represents the animal limens and the broken the human limens.

of the levels expressed logarithmically in Figure 11. It is seen that the line for the animals (the solid line) is at all levels higher than that for the human subjects (the dotted line). Their characteristic shapes, too, are different. While they both turn up at level I the slope is much sharper for the animal subjects. The relative threshold over the four levels is much more nearly constant for the human subjects than for the animals. With the relatively rough instrument used in this experiment it might be expected, from the accepted results of König and Brodhun, Aubert, and others, brought together by Hecht (7, p. 761), that the relative threshold for the human subjects would be more

or less constant over this range. That it is not absolutely constant for man, however, has been shown by the work of Lowry (7, p. 762), who measured the difference threshold with precision and found it to be continuously changing over this brightness range. Again, the monkey relative threshold seems to approach constancy on the three upper levels. The results as they stand seem to indicate that the monkey brightness vision is not nearly as good at the lower brightness of level I as at the other levels.

There are, however, certain possibilities that the limens as determined for Brightness Level I for the animal subjects are somewhat in error. As has already been noted, the variability among subjects at this level was greater than at any other. The experimental routine of testing was slightly different from that of the other levels (one experimental period each day during the summer months). Dido's limen for level I is much more nearly the same order of magnitude as her limens at the other levels than is the case with the other animals. Again, the state of adaptation of the subjects' eyes at level I was in all cases above that which would be obtained from long regard of the standard brightness of that level. It was indicated in the section on calibration that the average brightness to which the subjects were exposed between trials was between the standard brightnesses of levels II and IV. Cobb (1) has found that the difference threshold in man is lowest when the eye is adapted to the standard of the test field, or is surrounded by a peripheral field of the same brightness as the test field. Thus, in this experiment

we would expect from the conditions of experimentation that the limens on level I would be somewhat higher than those for the other levels. Also, with the animal subjects, who were observed on occasion to look directly at the dome light in their reaction cage, the adaptation might be to a very bright light, much brighter than any regarded by the human subjects, who never looked at the dome light above the screen used with them. We cannot, therefore, be entirely sure that the turning-up which we found in our human sensitivity curve, and to a greater extent in the animal curve, does not in part incorporate an experimental error. Though the absolute value of the limens which we determined for the monkeys on level I may be somewhat in error, the great difference for these and the ones for the human subjects at this level makes us inclined to believe that the difference is due in part at least to a real difference in the shapes of the monkey and human brightness-sensitivity curves.

Concerning the absolute differences between the limens of the monkey and the human subjects, we hesitate to conclude that we have determined the magnitude or even the direction of these differences. As has already been suggested, we do not know about relative general and specific practice which the two types of subjects may have had at brightness discrimination. Relative age and maturity status is about the only basis we have on which to match the subjects. Obviously, this criterion does not match them as to practice. Probably the human subjects had the most practice in everyday life, in reading, writing, and various

kinds of fine visual work; certainly the adult subjects had. Again, we cannot compare the general set of the monkey and human subjects in the apparatus. The two adults, who were graduate students in psychology, were acquainted with the general aim of the investigation. The child was apparently made to feel at ease in the situation, and was told the general nature of the experiment. We obviously cannot say much about the general "set" of the monkeys. They seemed to work well and not to show undue apprehension except in the case of Ulysses, as noted in the Protocols. Whether they worked to capacity or not we cannot tell. Time scores and consistency of performance indicate a relative constancy of motivation and of other experimental conditions. Also the difference in the type of response used by the animals and the adult human subjects definitely distinguishes our subjects. Had the adults been required to make a motor rather than a verbal response, as was the child, the limens so obtained might have been slightly different.

Despite the obvious difficulties with animal and human comparisons in this situation we feel that the relatively small differences in the limens for the two groups of subjects on the three upper levels of this experiment indicate that the monkey and human brightness-discrimination capacities are not greatly different over this range of brightnesses (0.7 to 50.0 millilamberts). At the lowest level (0.08 millilamberts) the difference appears to be real. It therefore seems justifiable to state that over the normal range of daylight vision the monkey can be presented with as critical brightness-discrimination problems as can the human subject.

IV

SUMMARY AND CONCLUSIONS

Limens for brightness discrimination at four absolute levels (0.08, 0.77, 7.1, and 55.3 millilamberts) with the light-adapted eye were determined for four young Rhesus monkeys. The animals were trained to reach toward the brighter window of a Yerkes-Watson light-discrimination apparatus for food. After preliminary training a modified method of limits was used at each level, with steps presenting brightness differences ranging from easy discriminability to equality. The percentages of correct choices made at each step on successive experimental periods were grouped and limens calculated by the quartile method (75-per-cent-correct choices). Discrimination times for three monkeys at two levels were grouped and analyzed. Two adult human subjects and one child of seven were tested under experimental conditions similar to those for the monkeys.

From the results of the experiment the following conclusions are drawn.

1. The average 75-per-cent limens for the monkeys in terms of fractional brightness difference (DI/I) were found to be: level I, 0.36±.026; level II, .10±.008; level III, 0.12±.009; level IV, 0.11±.008.

2. The average limens found with the two adult human subjects were slightly lower than those of the monkeys on the three upper levels, and a great deal lower on the lowest level. Differences in the amount of practice of the animal and human subjects, the difference in the directions that could be given, and dif-

ferences in the type of response make it impossible to say how valid are the absolute differences between the limens of the two types of subjects.

3. A real difference in the characteristic shapes of the monkey and human brightness-sensitivity curves is indicated, since the relative difference threshold was found to be much more nearly constant over all the levels with the human subjects than with the monkeys.

4. The results indicate that the Rhesus monkey may be presented with stimuli which necessitate a brightness discrimination about as fine as that which could be made by man over the normal range of day-light vision.

5. The results with human subjects show that the Yerkes-Watson light-discrimination apparatus is a more reliable and valid instrument for studying brightness vision than has been indicated by the work of Graham and Nafe.

6. Discrimination time for the monkeys has been found to increase in a more or less regular trend with a decrease in the brightness difference of the stimuli.

7. Wrong responses, in general, required slightly longer discrimination times than did correct ones.

8. Three modifications of the method of limits, a descending series with one response on each step, a descending followed by an ascending series with one response on each step, and a descending series with five responses on each step, all gave about the same limens (disregarding practice effects), with the two monkeys used in preliminary experimentation. The last method was the most convenient for use in the experiment proper.

References

1. Cobb, P. W. Vision and brightness of surroundings. *Trans. Illum. Eng. Soc.*, 1916, 11, 372-398.
2. Elder, J. H. Auditory acuity in the chimpanzee. *J. Comp. Psychol.*, 1934, 17, 157-183.
3. Fjeld, H. A. Limits of learning ability in Rhesus monkeys. *Genet. Psychol. Monog.*, 1934, 15, Nos. 5-6, 369-537.
4. Graham, C. H., & Nafe, J. P. Human intensity discrimination with the Yerkes-Watson apparatus. *J. Genet. Psychol.*, 1930, 37, 220-231.
5. Haggerty, M. E. Imitation in monkeys. *J. Comp. Psychol. & Neur.*, 1909, 20, 337-345.
6. Hartman, C. G., & Straus, W. L., Jr. [Ed.] The anatomy of the Rhesus monkey. Baltimore: Williams & Wilkins, 1933. Pp. 383.
7. Hecht, S. The nature of the photoreceptor process. In *A handbook of general experimental psychology*, ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press. London: Oxford Univ. Press, 1934. Pp. 704-828.
8. Jastrow, J. A critique of the psycho-physic methods. *Amer. J. Psychol.*, 1887-8, 1, 271-309.
9. Johnson, H. M. Visual pattern discrimination in the vertebrates. I. Problems and methods. *J. Anim. Behav.*, 1914, 4, 319-339.
10. ———. Visual pattern discrimination in the vertebrates: II. Comparative visual acuity in the dog, the monkey, and the chick. *J. Anim. Behav.*, 1914, 4, 340-361.
11. Kinneman, A. J. The mental life of two *Macacus rhesus* monkeys in captivity. I. *Amer. J. Psychol.*, 1902, 13, 98-148.
12. Klüver, H. Behavior mechanisms in monkeys. Chicago: Univ. Chicago Press, 1933. Pp. xvii+387.
13. Neet, C. C. Visual pattern discrimination in the Macacus Rhesus monkey. *J. Genet. Psychol.*, 1933, 43, 163-196.
14. Shepherd, W. T. Some mental processes in the Rhesus monkey. *Psychol. Monog.*, 1910, 12, No. 52. Pp. 66.

15. SPENCE, K. W. A study of visual acuity and its relation to brightness in chimpanzee and man. *J. Comp. Psychol.*, 1934, 18, 333-362.

16. WARDEN, C. J., JENKINS, T. N., & WARNER, L. H. Introduction to comparative psychology. New York: Ronald, 1934. Pp. x+581.

17. ———. Comparative psychology. New York: Ronald. (in preparation.)

18. WATSON, J. B. Behavior, an introduction to comparative psychology. New York: Holt, 1914. Pp. xii+439.

19. WOLLARD, H. H. The differentiation of the retina in the primates. *Proc. Zool. Soc. London*, 1927, 1, 1-17.

20. YERKES, R. M., & WATSON, J. B. Methods of testing vision in animals. *Behav. Monog.*, 1911, 1, No. 2, 1-28.

LA DISCRIMINATION DE BRILLANCE CHEZ LE SINGE RHÉSUS

(*Macacus Mulattus*)

(Résumé)

Les seuils de discrimination de brillance à quatre niveaux absolus (0,08, 0,77, 7,1, et 55,3 millilamberts), avec l'oeil adapté à la lumière, ont été déterminés pour quatre jeunes singes rhésus (*Macacus mulattus*). On a entraîné les animaux à tendre la main vers la nourriture à travers l'ouverture à volet de la cage de réaction, à travers l'une de deux portes de verre clair d'une "boîte de stimulus", vers la fenêtre plus brillante de verre opalin d'un appareil de lumière Yerkes-Watson. Après l'entraînement préliminaire on a testé les animaux avec diverses modifications de la méthode de limites. Une méthode où l'on a présenté cinq choix à chacune d'une série de situations de stimulus, causant une diminution progressive de la différence de brillance d'une discrimination facile jusqu'à l'égalité, s'est montrée la plus satisfaisante, en donnant les seuils les moins élevés et en maintenant une haute motivation pendant chaque période. On a présenté à chaque niveau de brillance une série contenant de huit à dix étapes de différence de brillance, de sorte que l'animal a fait de quarante à cinquante réponses de discrimination à chaque période expérimentale. On a testé les singes pendant trente à cinquante périodes expérimentales à chaque niveau de brillance pour éliminer les effets de l'exercice. On a déterminé les seuils (la différence fractionnaire de brillance où 75% des choix corrects a été fait) d'après les résultats des vingt dernières périodes expérimentales, c'est-à-dire, des 100 dernières réponses de discrimination faites à chaque étape de la différence de stimulus à chaque niveau. On a noté les temps de discrimination. On a employé des pommes et des raisins secs commes stimulants. On n'a employé aucune punition que le manque d'obtenir la récompense. On a testé deux adultes humains et un enfant

dans de semblable conditions expérimentales. Ils ont fait des réponses verbales (adultes) ou manuelles (enfant) et on leur a dit si leurs choix étaient justes ou faux. On a donné moins d'épreuves aux sujets humains qu'aux sujets animaux. On s'est servi de contrôles pour assurer l'élimination des réponses aux repères secondaires. On a fait des tests où l'on a renversé la brillance et l'intensité lumineuse des stimuli. Les seuils pour les singes dans les différences fractionnaires de brillance (DI/I) pour les quatre niveaux ont été Niveau I, 0,36±0,026; Niveau II, 0,10±0,008; Niveau III, 0,12±0,009; Niveau IV, 0,11±0,008; et pour les sujets humains: Niveau I, 0,015±0,013; Niveau II, 0,061±0,008; Niveau III, 0,070±0,011; Niveau IV, 0,144±0,022 (un sujet human au Niveau IV). Les courbes caractéristiques de la sensitivité différentielle de brillance pour les sujets animaux et les humains semblent être différentes. Tandis que les deux sont presque constantes pour les trois niveaux supérieurs, le mouvement en haut au niveau le plus bas est relativement beaucoup plus vertical pour le singe que pour l'homme, ce qui indique une plus grande efficacité pour l'oeil humain dans les illuminations inférieures. On a constaté que le temps de discrimination est devenu plus long dans une tendance assez régulière avec la diminution de la différence de brillance entre les stimuli. Plus de mouvements de comparaison ont été faits par les animaux aux plus petites différences de clarté. On a constaté que les réponses fausses ont exigé des temps de discrimination un peu plus longs que les réponses justes. On a standardisé une méthode cohérente pour la détermination des seuils visuels chez les primates sub-humains.

CRAWFORD

HELLIGKEITSUNTERSCHEIDUNG BEIM RHESUSAFFEN

(*Macacus mulattus*)

(Referat)

Die Unterscheidungsschwellen für vier absolute Schichten (0,08, 0,77, 7,1, und 55,3 Millilamberts) beim lichtangepassten Auge wurden für vier junge Rhesusaffen (*Macacus mulattus*) bestimmt. Die Tiere wurden trainiert, nach Futter durch ein Fensterchen im Reaktionskäfig zu greifen; durch eine von zwei klaren Glastüren eines "Triebkastens" (incentive-box) zu greifen; nach dem hellerem Opalglasfenster eines Yerkes-Watson-Apparates zu reichen. Nach einer Vorübung wurden die Tiere mit verschiedenen Veränderungen der Methode der Grenzen geprüft. Eine Methode, in der fünf Wahlen bei jeder einer Reihe von Reizstellungen dargeboten wurden, die eine forschreitende Abnahme des Helligkeitsunterschiedes von leichter Unterscheldbarkeit bis zur Gleichheit verursachten, erwies sich als die befriedigendste Methode zur Darbietung der niedrigsten Schwellen und zur Behauptung der hohen Motivierung durch jede Periode hindurch. Eine Reihe von acht bis zehn Stufen des Helligkeitsunterschiedes wurde bei jeder Helligkeitsschicht dargeboten, so dass von vierzig bis fünfzig Unterscheidungsantworten vom Tier Während jeder experimentellen Periode gemacht wurden. Die Affen wurden während dreissig bis fünfzig experimentellen Perioden bei jeder Helligkeitsschicht geprüft, um die Übungswirkungen auszuschalten. Die Schwellen (der Bruchhelligkeitsunterschied, bei dem 75 Prozent richtige Antworten gemacht wurden) wurden von den Ergebnissen für die letzten zwanzig experimentellen Perioden, d.h. von den

letzten 100 Unterscheidungsantworten, die bei jeder Stufe des Reizunterschiedes auf jeder Schicht gemacht wurden, bestimmt. Unterscheidungszeiten wurden aufgeschrieben. Äpfel and Rosinen wurden als Antriebe gebraucht. Keine Bestrafung ausser der Nichterhaltung der Belohnung wurde gebraucht. Zwei erwachsene Menschen und ein Kind wurden unter ähnlichen experimentellen Umständen geprüft. Sie machten Verbalantworten (Erwachsene) oder Handantworten (Kind) und ihnen wurde gesagt, ob ihre Wahlen richtig oder falsch wären. Weniger Proben wurden den menschlichen Vpn. als den Versuchstieren gegeben. Kontrollen wurden gebraucht, um de Ausschaltung der Antworten auf Nebenreize zu versichern. Prüfungen wurden ausgeführt, in denen die Helligkeits- und Lichtintensität der Reize umgekehrt wurden. Die Schwellen fur die Affen in Bruchhelligkeitsunterschieden (DI/I) für die vier Schichten waren: Schicht I, 0,36±0,026; Schicht II, 0,10±0,008; Schicht III, 0,12±0,009; Schicht IV, 0,11±0,008; und für die menschlichen Vpn.: Schicht I, 0,105±0,011; Schicht II, 0,061±0,008; Schicht III 0,070±0,011; Schicht IV, 0,144±0,022 (eine Versuchsperson auf Schicht IV). Die charakteristischen Kurven der Differentialhelligkeitsempfindlichkeit für die Versuchstiere und Versuchspersonen scheinen verschieden zu sein. Während beide beinahe unveränderlich für die drei Oberschichten sind, ist der Aufstieg bei der niedrigsten verhältnismässig viel steiler beim Affen als beim Menschen, was auf eine grössere Wirksamkeit für das menschliche Auge bei niedrigeren Helligkeiten hinweist. Die Unterscheidungszeit zeigte eine Zunahme der gewöhnlichen Richtung bei einem abnehmenden Helligkeitsunterschied zwischen Reizen auf. Mehr Vergleichsbewegungen wurden von den Tieren bei den kleineren Helligkeitsunterschieden gemacht. Falsche Antworten verlangten ein wenig längere Unterscheidungszeiten als die richtigen. Eine zuverlässige Methode für die Bestimmung der Gesichtsschwellen bei den Menschenaffen wurde standardisiert.

CRAWFORD

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

June, 1935
Volume XVII, No. 3

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

June, 1935
Volume XVII, No. 3

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

THE LIMITS OF LEARNING ABILITY IN CEBUS MONKEYS

From the Animal Laboratory of the Department of Psychology, Columbia University

By

ADOLPH MEYER KOCH

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

ACKNOWLEDGMENTS

The writer wishes to express his deep appreciation to Professor C. J. Warden of the Department of Psychology, Columbia University, for the supervision of the experiment and for much kindly criticism in the preparation of the manuscript. He is grateful to Mr. S. Dobrin for aid in computing the statistical indices and for rechecking the data.

ADOLPH MEYER KOCH

COLUMBIA UNIVERSITY
NEW YORK CITY

CONTENTS

I

INTRODUCTION

The behavior of monkeys has been investigated from numerous angles during the past few decades. A survey of the work in which problem boxes, simple tools, and multiple-choice devices were employed has been made recently by Fjeld (8). Since the present problem lies within this field of experimentation, a review of the literature will be unnecessary in this connection. It may be pointed out, however, that most investigators have used only from one to three monkeys in their work. Such factors as age, diet and care, preliminary taming, and conditions of testing were not sufficiently well standardized. Moreover, the rhesus monkey has been utilized in behavior studies much more extensively than the cebus monkey. This is probably due to the fact that the former is more readily available and has the reputation for being more hardy than the cebus when kept under ordinary laboratory conditions.

The investigation of the rhesus monkey by Fjeld (8) marked a definite forward step in the application of standard methods to the lower primates. She used the Jenkins problem box in which a graded series of tasks can be presented to the subject with the aim of determining the limits of learning capacity on this device. She also attempted to standardize age, diet (kind and amount of food at regular intervals), degree of hunger, method of handling and taming, and experimental conditions (incentives, illumination, elimination of distractions from other animals and experi-

menter, length and distribution of trials). She was also the first to make use of a sufficient number of monkeys (17 rhesus) to secure group indices on which the usual statistical comparisons could be based.

The present problem may be regarded as an extension of the method used by Fjeld to the cebus monkey. The general opinion seems to be that the cebus is less intelligent than the rhesus. There is little or no behavior evidence for this notion since comparable tests of consequence have not been carried out on those types. Tilney (22) seems to feel that the rhesus possesses an intelligence level considerably higher than that ascribed by Thorndike (20), in his earlier work, to New World monkeys. Bierens de Haan (1) and Klüver (12), on the other hand, report that the rhesus is far inferior to the cebus in the use of instruments. As a matter of fact not enough work has been done on the two types to make a fair comparison possible. The aim of the present experiment was to determine the limits of learning in the cebus monkey under the same conditions used by Fjeld in testing rhesus, so that limits of learning in the two types could be directly compared. It was also hoped that type differences in mode of attack on the problem would be brought out, if any exist.

It may be of interest to note, in passing, certain structural characteristics which distinguish the cebus monkey from the rhesus. The cebus monkey has no cheek pouches, possesses a prehensile tail, and has 36 instead of 32 teeth. According to Schultz (9), the five-

year-old male rhesus weighs 4310 grams, and Tilney (22) states the brain weight as 126 grams. The corresponding figures for the cebus are 1290 grams and 69.50 grams, as given by Leche (13). The body-brain ratio in terms of weight is thus 34.2 for the rhesus and 18.5 for the cebus. It thus appears that the cebus has a somewhat smaller brain than the rhesus, but a much larger brain per unit of body weight. In general, however, the New World monkeys are regarded as structurally more primitive than the Old World monkeys. While this is true, the inference that they are less intelligent is hardly warranted on the basis of structural differences alone. No attempt will be made in the presnet study to relate our behavior findings to structural differences in the two types of monkeys.

II

METHOD AND PROCEDURE

The task of taking the data in the present experiment extended from July 1933 to September 1934. Since our primary aim was to compare the cebus with the rhesus monkeys the method and procedure employed in our work was essentially the same as that of Fjeld (8). Our monkeys were kept in the same living quarters as the rhesus had been and tested in the same experimental room. Each monkey was kept in a separate cage 35.5 inches deep, 27.5 inches wide, and 30.5 inches high. The cages were arranged in a quadrangle and the enclosed court provided a suitable place for exercise. In as much as the method and procedure have been described in detail by Fjeld, it will be necessary to indicate here only the more important factors involved.

Apparatus

The apparatus used was a large model of the Jenkins Problem Box, a drawing of which is shown in Figure 1. The cage is constructed of a steel framework of heavy wire mesh, and rests upon a maple base painted battleship gray. The main cage was 7 feet in diameter and all the dimensions were in proportion. The three reaction plates were located in the floor as shown in Figure 2. They were set flush with the floor, were wired for electric shock, and were adjusted to the weight of the animal. The inner door of the entrance compartment was operated by hand admitting the an-

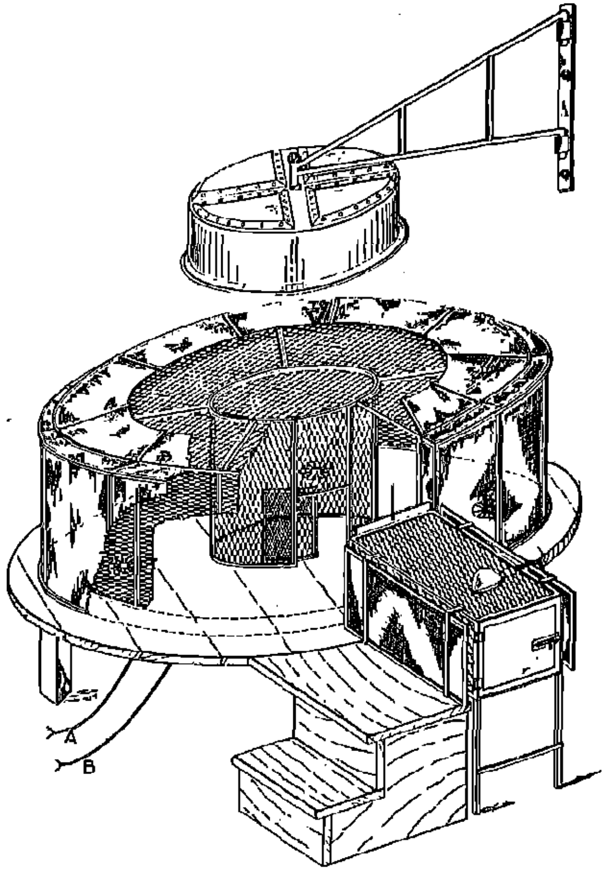


FIGURE 1

JENKINS PROBLEM BOX

Parts of the one-way screen and cage have been removed in order to make the interior of the apparatus more visible. The three plates in the floor are wired for shock, the current being supplied through *A* from the Jenkins Stimulation Apparatus (not shown). The electrical system (not shown) operates through *B* to open the door of the incentive cage when a given problem has been solved. The cord leading up from the door of the entrance compartment runs over a pulley and is operated by the experimenter some distance away. The large central illumination chamber is shown pushed back from the apparatus. (By courtesy of Professor C. J. Warden.)

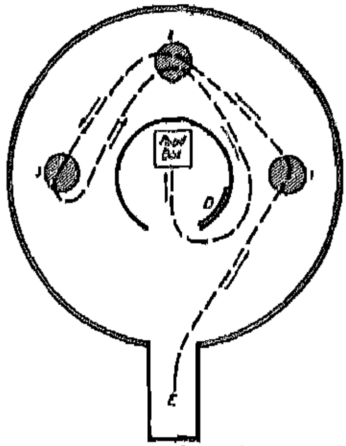


FIGURE 2

FLOOR PLAN OF THE JENKINS PROBLEM BOX

The broken line indicates a perfect response in Step IV. The animal starting at *E* is required to step on the plates in the order 1, 2, 3 and reverse to 2, to open the door *D*, and obtain access to the food. (By courtesy of Professor C. J. Warden.)

imal into the experimental cage. The door of the incentive compartment was held shut by an electromagnetic device and was so constructed that it opened by force of gravity when released and came to rest by its own inertia. This door was thus practically noiseless.

The large circular cage was lighted by four 25-watt and eight 40-watt bulbs set in a large reflector 38 inches in diameter by 13 1-4 inches deep. The reflector was larger in diameter than the inner cage so as to eliminate shadows from the framework and walls of the latter. The light was diffused through sheets of opal and plate glass into the apparatus below. The entrance compartment was lighted by a 10-watt bulb

set in an ordinary reading-lamp reflector closed at the bottom with a sheet of opal glass. Tests were conducted in a room otherwise dark and the experimenter was eliminated by means of a one-way light screen.

The switchboard and relays by means of which the operation of the inner door was controlled was located on the experimenter's control table at the right of the apparatus. This switchboard could be set so that the inner door opened automatically when any desired number of plates had been depressed. The depressing of the plates could be observed either directly or by observing the small signal lights on the switchboard. The experimental records were taken at the control table, the necessary light for this being supplied by a 10-watt pilot light partly concealed in a dark box. A 16-inch electric fan was run constantly at medium speed to serve as a noise screen.

SUBJECTS

Six cebus monkeys were used in the experiment, five of these being white-throated capuchins and one of the brown variety. All of them were males. The exact age of the animals is not known since they were obtained from a local dealer. They were all prepubertal, however, and probably ranged in age from about 24 to 30 months. Very little is known regarding the growth curve in cebus monkeys. The weights indicated in the protocols were obtained shortly before the experiment proper was begun.

Protocols

Monkey A. Weight, 34 ounces. Limit, 6 steps.

This monkey was in excellent physical condition and adapted very readily to the training situation. He was very tame and would eat out of my hand within a few days after being brought into the laboratory. He worked rapidly and consistently most of the time in the problem box. On the seventh step he appeared to lose interest in the problem and kept repeating the same errors trial after trial.

Monkey B. Weight, 35 ounces. Limit, 14 steps.

This monkey was in good physical condition throughout the experimental period. He tamed less readily than the other monkeys and was always shy, but became well adapted to the laboratory routine in time. As a rule his reactions were very rapid and definite. Sometimes he would complete three or four trials in good order and then would dash around the apparatus paying very little heed to the plates. On some days he worked well throughout the series. This inconsistence of response was present throughout the experimental series.

Monkey C. Weight, 37 ounces. Limit, 15 steps.

This animal appeared in perfect health but was less playful than the other monkeys. He was very tame and reacted consistently throughout the experiment.

Monkey D. Weight, 39 ounces. Limit, 5 steps.

The physical condition of this animal was excellent, but he showed a marked tendency to be inactive. During the preliminary period he seemed inclined to sit near the door instead of exploring the box. In the early trials of the test period proper he often climbed upon the wire in quadrant *A*, chewing at the entrance door or clawing it with his hands. It was thought at first that he was not well motivated, hence a bit of peanut, for which he showed a distinct preference, was added to the incentive (apple and raisin). This had little effect as he was easily discouraged and spent much time in sitting about. Increasing the

time per trial from three to five minutes was also ineffective. The animal was quite tame and showed no hesitancy in entering the apparatus.

Monkey E. Weight, 42 ounces. Limit, 11 steps.

This animal was of the brown variety. It showed a tendency to get diarrhea but this was remedied by reducing the amount of green vegetable in the diet. He tamed easily and was inclined to be playful. His behavior on the first trial was somewhat peculiar. He tended to run across a plate and then touch it again with his hind legs by backing up to it. This peculiarity disappeared before the first step had been learned and did not reappear thereafter.

Monkey F. Weight, 32 ounces. Limit, 8 steps.

This subject was in good physical condition throughout and tamed very readily. He had a peculiar habit of keeping his fingers in his mouth much of the time and would attempt to lick the experimenter's fingers when possible. His movements in the apparatus were slow and he was easily discouraged. If he did not secure the food after slight effort he would often sit and wait until readmitted to the entrance compartment for the next trial. This behavior was especially noticeable during the early trials of each step.

Since the experiment was to last over a considerable period of time the care and feeding of the animals was a matter of considerable importance. After consulting various authorities at the zoological park it was decided to follow the feeding schedule which had been worked out at the Columbia laboratory for the rhesus monkey. This schedule, which provides this somewhat varied diet from day to day, is given below:

	Morning	Afternoon	Evening
Sunday:	Milk and eggs	Wholewheat bread and raw carrot, also two peanuts	One-half banana[1]

[1]One sixteenth of a teaspoon of powdered Adex tablets mixed with bone meal was spread over each portion of banana to insure the sufficiency of vitamins A and D.

Monday:	Milk and eggs	Boiled rice and celery	One-half banana
Tuesday:	Milk and eggs	Boiled potato and lettuce	One-half banana
Wednesday:	Milk and eggs	Wholewheat bread, lettuce, and two peanuts	One-half banana
Thursday:	Milk and eggs	Boiled rice and celery	One-half banana
Friday:	Milk and eggs	Wholewheat bread and lettuce	One-half banana
Saturday:	Milk and eggs	Boiled sweet potato and raw carrot	One-half banana

The general health of the animals was excellent throughout the experiment. The brown cebus (subject E) developed diarrhea on occasions so that a strict regulation of the amount of greens allowed this animal was found necessary. The daily schedule followed throughout the experiment was as follows:

Early morning. Monkeys released together in central enclosure for exercise and exposure to sunlamp (General Electric Sunlamp, Model B, vertically adjustable, Mazda lamp, type S-1 bulb).
8:00 to 8:30 Morning meal.
9:00 to 11:00. Cages cleaned and supplied with water.
11:00 to 3:00 P. M. Experimental period (this varied somewhat from step to step).
After experimental period. Afternoon meal.
6:00. Evening meal.

A few facts should perhaps be noted regarding the adjustment of the animals to the laboratory conditions. Upon arrival at the laboratory a leather belt was placed on each monkey. Attached to each belt by a ring was a light chain approximately six inches long

with a ring at the end. Onto this ring was snapped a longer chain which served as a leash. After being in the laboratory for three days each monkey was taken out on the leash, fed sunflower seeds, and otherwise subjected to the taming process. The animals were sufficiently well tamed to begin experimental work in from one to two weeks. The period of adjustment was very much shorter than that found necessary by Fjeld for the rhesus monkeys. An attempt was made to bring all the animals up to the same degree of tameness. In no case was a monkey treated as a pet, the taming process being qualitatively the same for all.

General Procedure

After being tamed each monkey was placed in the entrance compartment for five minutes daily for a week or so. Raisins and bits of apple were dropped in to him during this period. After becoming adapted to the entrance compartment, the animal was introduced each day into the incentive compartment by means of a tunnel. This tunnel connected the entrance compartment and the incentive compartment and kept the animal from getting out into the main cage. Small bits of raisin and apple were placed in the incentive compartment and the animal allowed to secure these upon release from the entrance compartment. Since the monkey must return to the entrance compartment after securing the incentive each time this response soon became firmly established. In case the return was not prompt a slight pull on the long cord attached

to the animal served as a cue to this response. It seemed likely that the flashing on of the light in the entrance compartment when the door was raised also served as a cue. A few trials a day over a period of from one to two weeks fixed the habit of returning to the entrance compartment without help from the experimenter. In the case of subject *E*, the long cord had to be retained through the experiment proper since the animal occasionally failed to return with sufficient promptness. As soon as an animal had learned this routine the regular experimental series was begun without the tunnel.

The procedure for these trials in the regular experiment was as follows:

1. The incentive compartment was baited, the large reflector lighted, and the room lights turned off.
2. The monkey was brought into the experimental room and placed in the entrance compartment.
3. The door into the reaction compartment was raised by the experimenter who was seated at the control table.
4. This door was lowered behind the animal as it left the entrance compartment.
5. After performance the animal secured the food in the incentive compartment and returned to the entrance compartment.
6. The apparatus was then reset for the next trial.

After testing for food preference a small raisin and a bit of fresh apple were adopted as the incentives. These foods were strictly excluded from the regular

diet of the animal. Since subject *D* showed a marked preference for peanuts, one-fourth of a peanut kernel was included in the bait for this animal. Motivation appeared to be at a high level throughout the experiment. Each animal was given three sunflower seeds after every five-trial series in order to keep the animals used to the experimenter.

In a few cases the monkey tended to exhibit stereotyped errors over a series of trials. Such errors could be broken up fairly readily by the use of delay as punishment. A delay of 15 seconds before allowing the animal to enter the incentive compartment after each stereotyped performance was usually effective in time. In such cases the animal was required, of course, to correct its performance.

During the first two weeks the monkeys were brought up gradually to the level of five trials per series, and this rate was continued thereafter. After all of the animals had completed one five-trial series, a second series was given in the same order. The norm of mastery was nine perfect trials out of ten. A perfect trial consisted in a direct response to the plate or plates in the proper order without irrelevant behavior. In problems involving several plates, it was necessary for the animal to indicate the end of the series either by hesitating after the last plate was touched off or by going directly to the food compartment. As soon as the required degree of mastery was obtained the animal was immediately transferred to the next step. Failures were recorded when the animal did not succeed in touching

off the necessary plates within three minutes. In such cases the animal was returned to the entrance compartment without reward. A corrected trial was recorded when the animal made errors during the trial but finally touched off the plates in the proper order within the three-minute limit. In some instances the errors were interspersed at different points within the true pattern. The animal was allowed access to the incentive after the completion of these corrected trials.

Two more or less distinct scores were taken: (*a*) learning scores—trials, errors, and minutes; (*b*) activity scores—number of plates depressed, quadrants traversed, number of plates touched per quadrant traversed, number of plates touched per minute, number of quadrants traversed per minute, and various types of errors. The term "quadrants" refers to a division of the floor into four sectors by imaginary lines which cut across the center of each plate. A description of various other kinds of behavior activities was included in the record and was useful in interpreting various behavior scores.

Special Procedure for the Various Problems

The experiment may be conveniently divided into two parts: (*a*) the basic problem, (*b*) the advanced problem. Since the procedure differed somewhat in the two problems it will be necessary to point out what the differences are:

Basic Problem. The task here consists of three steps. In Step I the monkeys were trained to touch

plate 1, the first plate to the right as they entered the reaction compartment (see Figure 2). In Step II, plates 1 and 2 must be depressed in order. In Step III, plates 1, 2, and 3 must be depressed in order. Each step was learned to nine perfect trials out of ten before the subject passed on to the next step. Electric shock was not used as punishment in connection with the basic problem.

Advanced Problem. The advanced problem involved a continuation of the pattern beyond the basic problem by graded steps. The reaction required in Step IV (first step of the advanced problem) is illustrated in Figure 2. As will be noted this requires a reversal of direction after stepping on plate 3, before stepping on plate 2 to complete the performance. Step V included the series of plates 1, 2, 3, and reversing direction to 2 and 1. An additional reversal of direction would be required, of course, on Steps VI, VIII, X, XII, and XIV. The subject was required to learn each step to nine perfect trials out of ten before being transferred to the next step. When an animal failed to reach this norm of mastery on a given step within 1000 trials, training was discontinued, and the preceding step was taken as the limit of learning for the animal. Since no animal required more than 562 trials to learn any step which was learned, the 1000-trial criterion for final failure would seem to be a generous one. As a matter of fact an animal seldom showed any tendency to improve during the latter half of the series.

In the advanced problem electric shock was used to

serve as a cue in learning the various steps. The strength of the shock was not very great and was given only when the animal stepped on the wrong plate. The shock was produced by an alternating current of 60 cycles, with a terminal pressure of 200 volts, an external resistance of 10,000,000 ohms in the circuit, and a current of 0.02 milliamperes. The apparatus which controlled the current was the Jenkins Stimulation Apparatus which is described in detail by Jenkins, Warner, and Warden (24). The above strength of shock, which was found suitable for the cebus monkey in our work, was much less than that utilized by Fjeld on the rhesus monkeys. This should be expected in view of the fact that the body weight of the cebus monkey is much less than that of the rhesus. Our observations led us to believe that the light shock used operated as a signal rather than as a punishment.

In general the advanced problem should be regarded as a continuation of the basic problem so far as the animal is concerned. Aside from greater complexity of pattern, the former involved a new element of a reversal of direction from time to time. On the side of procedure, the chief element of novelty in the advanced problem consisted in the use of shock for errors. It should be pointed out, perhaps, that the various steps on the apparatus varied greatly in difficulty. Any given pattern, therefore, does not represent a perfectly graded series of steps in terms of ease of mastery.

III

THE BASIC PROBLEM

The data secured in this experiment have been divided for the sake of convenience into two classes, learning data and activity data. Under learning data are included trial, error, and time scores. Under activity data are included all other scores representing a more detailed analysis of behavior.

LEARNING SCORES

The individual learning scores for the six monkeys in the three steps of the basic problem are presented in

TABLE 1

RESULTS OF BASIC PROBLEM SHOWING INDIVIDUAL LEARNING SCORES

Animals	Trials			Errors,	Minutes,
	Total	Failure	Perfect	total	total
			Step I		
Males					
A	80	3	16	351	33.75
B	243	5	52	1022	70.12
C	47	8	12	168	32.93
D	84	20	18	403	97.22
E	327	6	75	921	91.77
F	42	5	16	73	23.07
			Step II		
A	8	0	2	53	3.25
B	71	3	1	355	20.05
C	82	0	5	144	15.15
D	41	0	25	53	5.82
E	129	0	7	516	29.07
F	135	24	32	540	124.12
			Step III		
A	34	0	10	89	10.35
B	37	1	1	107	9.83
C	91	0	4	123	14.67
D	83	0	31	200	21.97
E	280	0	59	1197	71.02
F	80	0	32	172	22.65

Table I. These scores represent the learning period up to, but not including the norm. Thus, the last eight or nine trials were dropped off, depending upon whether there were nine or less perfect trials in succession. The total number of failure trials and the total number of perfect trials for each subject are shown in the table, as well as the total number of errors and minutes. As mentioned previously, a failure trial is one in which the animal fails to touch off the necessary plate or plates in the allotted time of three minutes, and thus does not secure the reward. A perfect trial is one in which the subject goes directly from the entrance compartment, touches off the plates in the required sequence, and enters the food compartment without hesitation or irrelevant behavior. Corrected trials have been defined in the previous section.

TABLE 2

SHOWING MEASURES OF CENTRAL TENDENCY AND VARIABILITY FOR TRIALS, ERRORS, AND MINUTES IN THE THREE STEPS OF THE BASIC PROBLEM

	Median	Average	Range	Quartile deviation	Average deviation	Standard deviation	Coefficient of variation
			Step I				
Trials	82.00	137.17	42-327	59.50	80.83	108.41	79.03
Errors	377.00	489.67	73-1022	270.75	292.33	358.99	73.31
Minutes	51.94	58.16	23.07-97.22	26.48	28.23	29.61	50.91
			Step II				
Trials	76.50	77.67	8-135	40.50	37.67	45.06	58.01
Errors	249.50	276.83	53-540	191.25	193.50	204.29	73.80
Minutes	17.60	32.91	3.25-124.12	10.01	24.84	41.69	126.68
			Step III				
Trials	81.50	100.83	34-280	25.75	50.50	83.15	82.47
Errors	147.50	314.67	89-1197	44.00	208.33	396.40	125.97
Minutes	18.32	25.08	9.83-71.02	6.11	13.47	21.15	84.33

The measures of central tendency and of variability are presented in Table 2. Due to the small number of subjects and the great variability, the median may be considered a better measure of central tendency than the average. The median is much smaller than the average for the total group in all cases.

The performance is characterized by a very wide range, whether measured in terms of number of trials, errors, or minutes. This shows that some of the monkeys learned very quickly and with relatively few errors, while others required a large number of trials and made very many errors. The extreme variability of the group is also shown by the other measures of variability. Due to the small number of animals and the spread of the scores, the quartile deviation is, perhaps, the most significant of these measures. The other measures, however, are presented in order to allow comparisons to be made with similar studies on other animals.

The high variability of the scores may be due in part to the nature of the experimental set-up and in part to genuine individual differences among the subjects. The experimental conditions were such as to bring about a decrease in the chance factor from step to step. In the initial attack upon Step I, for example, it was largely a matter of chance whether or not the monkey touched off the first plate within the time limit imposed. As shown in Table 1, there were 47 failure trials on Step I, 27 on Step II, and only 1 on Step III. These figures suggest that the amount of variability

may have been due in part to the difference in the weighting of the chance factor from one stage of the learning process to another.

The question arose as to whether the high variability may not have been the result of using a rigorous norm of mastery—nine perfect trials out of ten. In order to settle this point, the data were computed for each step on the basis of two lower norms: four perfect trials out of five (norm 2), and one perfect trial (norm 3). In making comparisons, it should be borne in mind that the higher norm was actually employed in the training. The figures for Step II and III, based upon the lower norms, do not show what would have happened if the animal had been pushed ahead as soon as these lower stages of fixation were actually reached. The averages and variability indices for the three norms on the several steps are given in Table 3. The averages are, of course, much lower when computed on the basis of the less rigorous norms of mastery. The influence of the norm factor on variability, however, is not very great in most cases. Nor is the amount of variability always less for the lower norms. It seems unlikely, therefore, that the use of a rigorous norm had any consistent influence upon the variability of the scores.

The high variability seems to be due primarily to diversity of characteristics obtaining among the few subjects used. According to Fjeld (8) the most important of these for the monkey in this problem box are the following: differences in age, in playfulness,

TABLE 3

Showing Measures of Central Tendency and Variability for Trials, Errors, and Minutes for the Three Norms in the Three Steps of the Basic Problem

Norms	Median	Average	Range	Quartile deviation	Average deviation	Standard deviation	Coefficient of variation
			Step I				
			Trials				
1	82	137.17	42-327	59.5	80.83	108.41	79.03
2	67	82.17	17-191	32.5	44.5	56.60	68.88
3	37	54	9-150	27.25	39.67	49.31	91.31
			Errors				
1	377	489.67	73-1022	270.75	292.33	358.99	73.31
2	369.5	420.33	73-845	228.5	234.67	284.76	67.75
3	283	347.5	48-721	210.5	230.83	254.00	73.09
			Minutes				
1	51.94	58.16	23.07-97.22	26.48	28.23	29.61	50.91
2	44.28	53.05	23.07-95.47	21.64	24.72	27.66	52.14
3	38.03	48.81	16.92-95.42	22.0	26.10	29.57	60.58
			Step II				
			Trials				
1	76.5	77.67	8-135	40.5	37.67	45.06	58.01
2	53	55.5	8-129	34.25	38.5	43.35	78.11
3	21	20	3-37	12.5	11.33	13.90	69.50
			Errors				
1	249.5	276.83	53-540	191.25	193.5	204.29	73.80
2	178.5	218.17	28-516	121.75	143.17	17.46	78.59
3	118	134.5	26-284	77.75	88.5	94.99	70.62
			Minutes				
1	17.6	32.91	3.25-124.12	10.01	24.84	41.69	126.68
2	17.6	20.16	1.28-52.13	11.15	13.54	17.15	85.07
3	8.03	14.68	.72-52.03	5.7	11.35	17.48	119.07
			Step III				
			Trials				
1	81.5	100.83	34-280	25.76	50.5	83.15	82.47
2	37	45.5	11-91	20.25	22.5	28.68	63.03
3	11	19.83	1-66	5.25	12.83	21.37	107.77
			Errors				
1	147.5	314.67	89-1197	44	208.33	396.4	125.97
2	115	141.17	15-379	35.25	68.83	112.76	79.88
3	68.5	57.5	15-92	32.5	26.5	31.4	54.61
			Minutes				
1	18.32	25.08	9.83-71.02	6.11	13.47	21.15	84.33
2	13.87	12.30	1.60-19.17	4.68	4.14	5.54	45.04
3	8.11	6.65	.58-11.35	4.16	3.35	4.15	62.41

in the formation of stereotyped errors of the position-habit type, in the occurrence of specific reaction tendencies, in the amount of transfer effect from step to step, and in genuine individual differences in learning ability.

Although our monkeys were presumably of about the same age, no definite check on this point could be made. They were all in the prepubertal stage of development but the precise age could not be vouched for by the importer. The weight data secured proved to be of little value because the growth curve for the cebus monkey has not been worked out as yet.

The degree of tameness may have been effective but not in any very important way. The cebus monkey is much more readily tamed than the rhesus monkey. The protocols indicate that some of our animals were tamer than others in spite of our effort to equalize this condition. None of them showed a definite conditioning against the apparatus as sometimes occurs in rhesus monkeys.

Some of the monkeys exhibited too much playfulness at times as indicated by an increase in time and error scores. This was true of subject *E* throughout the tests on the basic problem, and of subject *A* on some of the advanced problems. In the latter case, the playfulness increased as the tests proceeded. Stereotyped behavior of the position-habit type was not observed. Specific reaction tendencies were exhibited by some of the subjects from time to time. This would seem to be indicated by the fact that there was little

or no consistency in the relative difficulty of the several steps from monkey to monkey. It is true that Step I was on the average more difficult than the other two steps. Nevertheless, Step I proved to be the easiest of all for subjects *C* and *F*. As Fjeld has pointed out, such differences in relative difficulty are probably due to specific reaction tendencies, peculiar to the individual, which influenced the rate of learning on one step more than on another. The influence of transfer-of-training effects will be discussed in a later section. In the last analysis, it would seem that the differences in score represent largely genuine differences in learning capacity. However, the lack of consistency in the scores of an animal indicates clearly that "ability to learn cannot be judged on the basis of one step alon ."

A modification of the Vincent method was used to represent trends in errors and time in the learning of the several steps. This method consists in taking the percentage of errors or minutes for each tenth of the total number of trials required to learn for each individual animal. These values were then combined so as to represent the corresponding group scores. These percentages are shown in Table 4. It will be seen that if these values were plotted in curves the form would be that of typical error and time learning curves. As might be expected, there are a few irregularities.

Activity Scores

In dealing with the activity scores I find it convenient to follow the classification of Fjeld by recogniz-

TABLE 4

SHOWING AVERAGE PERCENTAGE SCORES IN ERRORS AND TIME FOR EACH TENTH OF THE LEARNING PERIOD IN THE THREE STEPS OF THE BASIC PROBLEM

	Step I		Step II		Step III	
Tenths	Errors	Time	Errors	Time	Errors	Time
1st	24.98	36.85	32.91	29.18	17.05	17.35
2nd	19.51	21.87	9.58	13.00	13.67	12.34
3rd	13.41	11.50	5.57	3.95	6.83	7.93
4th	11.40	9.06	6.04	3.68	9.22	10.28
5th	8.51	6.46	10.29	13.97	7.41	7.16
6th	7.66	4.89	10.83	14.75	8.69	8.00
7th	3.68	2.75	6.40	5.75	8.69	9.69
8th	3.84	2.24	7.78	7.46	10.54	10.01
9th	3.00	1.67	5.99	4.89	9.27	8.71
10th	3.95	2.71	4.61	3.37	8.63	8.53
Totals	100.00	100.00	100.00	100.00	100.00	100.00

TABLE 5

RESULTS OF BASIC PROBLEM SHOWING INDIVIDUAL ACTIVITY SCORES

Animals	Plates depressed	Quadrants traversed	Quadrants per plate	Plates per minute	Quadrants per minute
			Step I		
A	223	521	2.34	6.61	15.44
B	480	2155	4.49	6.85	30.75
C	69	334	4.84	2.10	10.14
D	77	995	12.92	.79	10.23
E	609	1793	2.94	6.64	19.54
F	46	127	2.76	1.99	5.50
			Step II		
A	48	66	1.38	14.77	20.31
B	264	552	2.09	13.17	27.53
C	197	258	1.31	13.00	17.03
D	105	133	1.27	18.04	22.85
E	581	942	1.62	19.99	32.40
F	364	955	2.62	2.93	7.69
			Step III		
A	143	188	1.31	13.82	18.16
B	149	207	1.39	15.16	21.06
C	285	305	1.07	19.43	20.79
D	347	534	1.54	15.79	24.31
E	1608	2104	1.31	22.64	29.63
F	295	366	1.24	13.02	16.16

ing three categories: (1) scores obtained directly by observation, (2) scores computed from the original time and activity records, and (3) scores obtained by analyzing errors into various classes. These three groups of scores will be discussed in order.

The first group includes the number of plates depressed and the number of quadrants traversed. The first of these indicates the amount of activity directed toward the plate, while the last is a measure of general activity. Since both scores are directly affected by the length of the learning period they are fairly closely related to the learning scores. Thus if an animal took a long time to learn he would likely depress more plates and traverse more quadrants than if he learned more quickly. The medians and averages indicate consistent increases in the number of plates depressed from step to step. This should be expected because of the increase in the number of plates involved in the solution of each succeeding step. An opposite tendency was shown, however, with reference to the median and average number of quadrants traversed. This would seem to indicate that extraneous activities decreased from step to step as the attack upon the plates became more and more direct. The measures of variability, Table 6, indicate marked individual differences in the number of plates touched and the number of quadrants traversed. These differences agree fairly closely with the corresponding differences in the learning scores (Table 2). Doubtless the same factors contributed to variations from individual to individual in both cases.

TABLE 6

SHOWING MEASURES OF CENTRAL TENDENCY AND VARIABILITY FOR NUMBER OF PLATES DEPRESSED, NUMBER OF QUADRANTS TRAVERSED, NUMBER OF QUADRANTS PER PLATE, NUMBER OF PLATES DEPRESSED PER MINUTE, AND NUMBER OF QUADRANTS TRAVERSED PER MINUTE

	Median	Average	Range	Quartile deviation	Average deviation	Standard deviation	Coefficient of variation
Step I							
Number of plates depressed	150.00	250.67	46-609	147.00	186.67	218.64	87.22
Number of quadrants traversed	758.00	987.50	127-2155	551.75	660.17	752.45	76.20
Number of quadrants per plate	3.72	5.05	2.34-12.92	1.06	2.37	3.64	72.08
Number of plates per minute	4.36	4.16	.79-6.85	2.62	2.54	2.57	61.78
Number of quadrants per minute	12.84	15.26	5.50-30.73	4.89	6.64	8.21	53.80
Step II							
Number of plates depressed	230.50	259.83	48-581	118.75	143.17	176.45	67.91
Number of quadrants traversed	405.00	484.33	66-955	323.75	352.00	361.76	74.69
Number of quadrants per plate	1.50	1.72	1.27-2.62	.29	.40	.49	28.49
Number of plates per minute	13.97	13.65	2.93-19.99	1.66	3.95	5.42	39.71
Number of quadrants per minute	21.58	21.30	7.69-32.40	6.42	6.29	7.84	36.81
Step III							
Number of plates depressed	290.00	471.17	143-1608	57.50	278.83	513.97	109.08
Number of quadrants traversed	335.50	617.33	188-2104	126.25	384.00	674.60	109.28
Number of quadrants per plate	1.31	1.31	1.07-1.54	.10	.10	.14	10.69
Number of plates per minute	15.48	16.64	13.02-22.64	2.10	2.64	3.36	20.19
Number of quadrants per minute	20.93	21.69	18.16-29.63	2.77	3.34	4.36	20.10

These factors have already been discussed in connection with the learning scores.

A modified Vincent method, as mentioned above, was used to bring out trends in number of plates touched and number of quadrants traversed. A gradual decrease in these types of behavior occurred for each of the several steps, as indicated in Table 7. It should

TABLE 7

SHOWING AVERAGE PERCENTAGE SCORES FOR NUMBER OF PLATES DEPRESSED AND NUMBER OF QUADRANTS TRAVERSED FOR EACH TENTH OF THE LEARNING PERIOD ON THE THREE STEPS OF THE BASIC PROBLEM

	Step I		Step II		Step III	
Tenths	Plates	Quadrants	Plates	Quadrants	Plates	Quadrants
1st	14.53	37.27	20.78	30.18	13.20	15.23
2nd	15.20	19.01	8.79	9.12	11.64	11.77
3rd	13.14	11.29	8.27	5.75	8.70	8.04
4th	9.95	8.17	7.89	5.40	9.23	9.21
5th	9.75	6.43	10.01	10.98	8.77	7.86
6th	9.22	4.84	10.65	11.35	9.16	8.64
7th	7.17	3.76	8.47	6.92	9.76	9.75
8th	7.56	3.17	9.43	8.12	10.26	10.74
9th	6.84	2.04	8.34	6.61	9.69	9.15
10th	6.64	4.02	7.37	5.57	9.59	9.61
Total	100.00	100.00	100.00	100.00	100.00	100.00

be remarked that the results of Fjeld on the rhesus monkey show an increase in plates touched on Step I. This would seem to mean that the cebus monkey learned to react to the plates on the floor rather early in their performance on Step I. This would mean that the problem gradually narrowed down to the correct plate in Step I as in the other two steps in the case of our monkeys.

The second group of scores (computed scores) in-

cludes the number of quadrants traversed per number of plates depressed, the number of plates depressed per minute, and the number of quadrants traversed per minute. The quadrant-plate ratio was obtained for each individual on a given step by dividing the total number of quadrants traversed by the total number of plates depressed by that animal. A perfect quadrant-plate ratio would be, of course, 1, regardless of the step involved. This ratio gives a fairly accurate measure of the proportion of the excess activity which was directed toward the plates. The other two computed scores—number of plates depressed per minute, and the number of quadrants traversed per minute—require no explanation. Both of these scores indicate the relative speed of activity during the performance.

These three types of computed scores are given in the last three columns of Table 5, and the measures of central tendency and of variability in Table 6. The measures of central tendency indicate a marked improvement in the quadrant-plate ratio from step to step, while the measures of variability show very striking decreases in differences within the group. There is also a corresponding decrease in the range from step to step. The wide range on Step I was probably due, in part at least, to the chance factor, which affects this step more than later steps, and in larger part to the difficulty in learning to react to plates. The measures of central tendency for speed-of-activity scores show large increases in the scores from step to step while the coefficients of variation show an evident

trend toward decreasing variability. The group as a whole then shows a speeding-up of plate and quadrant activity accompanied at the same time by a decrease in the variability with each succeeding step. The data for the speed-of-activity curves on the three steps are shown in Table 8. These trends were obtained by the

TABLE 8

Showing Average Number of Plates per Minute and Quadrants per Minute for Each Tenth of the Learning Period on the Three Steps of the Basic Problem

	Step I		Step II		Step III	
Tenths	Plates per minute	Quadrants per minute	Plates per minute	Quadrants per minute	Plates per minute	Quadrants per minute
1st	1.71	17.21	5.62	15.22	14.28	21.59
2nd	3.01	14.79	5.34	10.32	17.72	23.48
3rd	4.94	16.70	16.54	21.41	20.62	24.98
4th	4.76	15.35	16.92	21.60	16.87	22.04
5th	6.53	16.93	5.66	11.57	23.01	26.99
6th	8.17	16.86	5.70	11.33	21.53	26.60
7th	11.27	23.28	11.61	17.68	18.93	24.76
8th	14.65	24.16	9.98	16.02	19.24	26.41
9th	17.76	20.86	13.47	19.90	20.88	25.84
10th	10.60	25.24	17.24	24.29	21.12	27.75

regular Vincent method. The curves for plates per minute and quadrants per minute roughly indicate ascending tendencies with drops interspersed here and there, for both curves. This further indicates the increase in activity from step to step as representative of the group.

The third general kind of activity scores refer to the various types of errors made in solving the problem. In the analysis of errors the following classification was used:

Omission of Plates. This type of error was checked when a plate was omitted, the preceding and succeeding plate in the series having been depressed.

Wrong Plates. This involves depressing a plate out of the proper sequence.

Wrong Direction. This includes either starting the trial in the wrong direction or approaching the right plate in the wrong direction.

Irregular Course. This involves zigzagging back and forth in the apparatus, and only one error was counted on a given trial.

Climbing. The error of climbing about the apparatus was checked only once in a trial.

Entrance Door. This included going to or waiting at the entrance door.

Inner Door. This included going to or working at the door of the incentive compartment.

Going Around. This refers to going around in the incentive compartment without touching plates.

Running Beyond. This involved a tendency to run around the box or to work at the plates instead of going directly to the incentive compartment when the series had been completed.

Miscellaneous. This class included such incidental reactions as reaching to the top of the cage instead of stepping on the plates, pulling at the string of the entrance door, and touching the plate too lightly to depress it.

An analysis of the errors into these several types is shown for the several steps in Table 9. The values

TABLE 9
TYPES OF ERRORS, SHOWING THE DISTRIBUTION OF EACH ON THE THREE BASIC STEPS

Types of errors	Step I	Step II	Step III
Omission of plates:			
Plate 1	.47	8.41	12.23
Plate 2	.00	.07	1.64
Plate 3	.00	.00	.56
Wrong plates:			
Plate 1	.00	24.11	17.09
Plate 2	14.13	5.56	17.09
Plate 3	11.15	5.89	15.01
Wrong direction	27.19	7.15	13.62
Irregular course	3.02	1.99	.69
Climbing	3.59	3.44	4.35
Entrance door	5.21	6.89	.44
Inner door	11.26	29.74	12.80
Going around:			
Counterclockwise	3.76	.66	.13
Clockwise	15.65	.60	.57
Running beyond	3.69	2.91	2.65
Miscellaneous	.88	2.58	1.13
Total	100.00	100.00	100.00

there given represent the percentage of each type of error when the total number of errors for each animal in learning a step is used as a base. The table indicates, generally, that in each succeeding step there is a tendency for the percentage of errors having to do with the plates to increase and the percentage of extraneous errors to decrease. Obviously, such a result might be expected since with each step one or more classes of errors concerned with plates are added. In examining the table for the three types of errors having the highest percentage frequencies for the various steps we notice that for Step I these frequencies are: wrong direction, going around (clockwise) and wrong plate 2. For Step II they are: inner door, wrong plate 1, and omission of plate 1; for Step III, wrong

plates 1 and 2 have the same frequency, and next in rank is wrong plate 3.

The distribution of the various types of errors for successive tenths of the learning process on each of the three steps was computed after the manner of Fjeld (8), Table 17. We have omitted our own tabular results in this connection, since the analysis does not appear to be very important. On Step I the following errors tended to increase in relative frequency as learning progressed: omission of plates, wrong plate, wrong direction, inner door, running beyond, and miscellaneous. On the other hand, all other types of errors show a relative decrease. On Step II the following errors showed a relative increase: omission of plate 1, wrong plate 2, wrong direction, and miscellaneous. A relative decrease occurred for wrong plate 1, irregular course, climbing, running to entrance door, and going around in both directions. The most striking tendency here exhibited was for the animal, on Step II, to work repeatedly at plate 1, which had been successful in the previous step. As this tendency diminished plate 1 was often omitted entirely, the animal running directly to plate 2 on entering the apparatus. On Step III, the trends were the same as those indicated by Fjeld for the rhesus monkey.

The relative frequencies of the several types of errors are doubtless closely related to the reaction tendencies of the monkeys. The various trends in types of errors, running through the learning process, reflect

the mode of attack in solving the different problems. Our analysis of errors is based on too few cases, however, to enable us to draw definite conclusions as to the nature of the reaction tendencies involved.

IV

ADVANCED PROBLEM

As noted above in section II, the advanced problem was simply a continuation of pattern formation beyond the basic problem to the limit of each animal. The problem began with Step IV, and the limits of our animals ranged from Step V to Step XV inclusive. Each animal was advanced from step to step, after reaching the norm of mastery in each case, until it reached the level at which it finally failed. The results will be presented in the following order: learning scores, and activity scores.

LEARNING SCORES

The individual learning scores of the six monkeys on the various steps of the advanced problem are presented in Table 10, and the measures of central tendency and of variability in Table 11. The number of animals included decreases step by step as final failures occur. The scores for the step on which an animal finally failed could not, of course, be included. It is evident that the group data become less reliable as the problem becomes more complex. The medians are somewhat smaller than the averages and may be taken as a better measure of central tendency than the averages. Since the number of cases is small, and the spread of the scores wide, the range and the quartile deviation are the most significant measures of variability.

TABLE 10

RESULTS OF ADVANCED PROBLEM SHOWING INDIVIDUAL LEARNING SCORES

Animals	Total	Trials, failure	Perfect	Errors, total	Minutes, total
		Step IV			
A	378	3	105	2609	186.77
B	56	0	6	184	12.03
C	360	54	50	4461	337.65
D	251	106	41	2006	386.18
E	80	0	9	506	31.28
F	151	16	38	1532	144.08
		Step V			
A	178	1	82	413	48.30
B	28	0	9	44	5.35
C	22	0	5	61	5.88
D	138	4	36	611	54.03
E	56	0	16	160	14.07
F	66	0	27	105	14.60
		Step VI			
A	216	0	64	845	69.27
B	150	15	23	846	83.60
C	89	1	12	291	34.67
D	Failed				
E	131	0	12	571	33.78
F	144	0	51	341	45.08
		Step VII			
A	Failed				
B	69	1	18	238	23.82
C	104	0	37	232	30.15
E	96	0	16	227	19.83
F	214	0	41	1144	95.28
		Step VIII			
B	216	35	64	1614	153.55
C	61	0	22	131	18.85
E	125	0	11	307	28.35
F	372	60	43	2345	344.52
		Step IX			
B	50	11	12	385	42.70
C	562	7	117	2749	231.47
E	114	0	24	252	25.95
F	Failed				
		Step X			
B	103	7	32	384	48.45
C	136	0	41	310	34.37
E	176	0	50	489	42.63

TABLE 10 (*Continued*)

Animals	Total	Trials, failure	Perfect	Errors, total	Minutes, total
		Step XI			
B	26	0	8	43	8.62
C	65	0	26	122	19.80
E	392	0	128	915	108.00
		Step XII			
B	177	5	66	470	74.68
C	58	0	19	162	18.15
E	Failed				
		Step XIII			
B	61	0	15	94	18.93
C	89	0	28	344	31.35
		Step XIV			
B	361	12	140	905	155.47
C	73	0	15	188	22.83
		Step XV			
B	Failed				
C	89	3	15	605	44.25
		Step XVI			
C	Failed				

The results definitely indicate that Step IV was very difficult. This is shown by the large number of trials, errors, and minutes required to learn it. Several of the other steps, however, appear to be quite difficult for certain animals which finally came to learn them. It is hazardous to compare in difficulty the stages beyond Step V since animals began to drop out at this point. It seems likely that Step IV is particularly difficult because it involved the first reversal of direction in pattern formation. As indicated in Tables 10 and 11, differences in individual performance were quite marked on all the steps of the advanced problem. The factors which make for differences here are apparently the same as those already discussed in connection with

TABLE 11

SHOWING MEASURES OF CENTRAL TENDENCY AND VARIABILITY FOR TRIALS, ERRORS, AND MINUTES, IN THE VARIOUS STEPS OF THE ADVANCED PROBLEM*

	Median	Average	Range	Quartile deviation	Average deviation	Standard deviation	Coefficient of variation
			Step IV (6 animals)				
Trials	201.00	212.67	56-378	118.75	117.00	126.79	59.62
Errors	1769.00	1883.00	184-4461	981.25	1230.00	1419.75	75.40
Minutes	165.43	183.00	12.03-386.18	120.28	120.55	140.82	76.95
			Step V (6 animals)				
Trials	61.00	81.33	22-178	38.50	46.00	57.44	70.63
Errors	132.50	232.33	44-611	117.00	162.33	209.09	90.00
Minutes	14.34	23.71	5.35-54.03	12.92	15.27	19.81	83.55
			Step VI (5 animals)				
Trials	144.00	146.00	89-216	24.50	29.20	40.97	28.06
Errors	571.00	578.80	291-846	236.50	211.80	237.36	41.01
Minutes	45.08	53.28	33.78-83.60	14.62	16.88	19.84	37.24
			Step VII (4 animals)				
Trials	100.00	120.75	69-214	17.50	38.25	55.38	45.86
Errors	235.00	460.25	227-1144	5.50	230.75	394.78	85.78
Minutes	26.99	42.27	19.83-95.28	5.16	20.45	30.83	72.94
			Step VIII (4 animals)				
Trials	170.50	193.50	61-372	77.50	100.50	116.85	60.39
Errors	960.50	1099.25	131-2345	741.50	880.25	919.51	83.65
Minutes	90.95	90.88	18.85-344.52	67.35	112.72	139.07	153.03
			Step IX (3 animals)				
Trials	114.00	242.00	50-562				
Errors	385.00	1128.67	252-2749				
Minutes	42.70	100.04	25.95-231.47				

*Since only one animal learned Step XV, the data for that step are not included in this table. See Table 10 for individual records.

TABLE 11 (*Continued*)

	Median	Average	Range	Quartile deviation	Average deviation	Standard deviation	Coefficient of variation
			Step X (3 animals)				
Trials	136.00	138.33	103-176				
Errors	384.00	394.33	310-489				
Minutes	42.63	41.82	34.37-48.45				
			Step XI (3 animals)				
Trials	65.00	161.00	26-392				
Errors	122.00	360.00	43-915				
Minutes	19.30	45.47	8.62-108.00				
			Step XII (2 animals)				
Trials	117.50	117.50	58-177				
Errors	316.00	316.00	162-470				
Minutes	46.42	46.42	18.15-74.68				
			Step XIII (2 animals)				
Trials	75.00	75.00	61-89				
Errors	219.00	219.00	94-344				
Minutes	25.14	25.14	18.93-31.35				
			Step XIV (2 animals)				
Trials	217.00	217.00	73-361				
Errors	546.50	346.50	188-905				
Minutes	89.15	89.15	22.83-155.47				

the basic problem. The most important factor, perhaps, is genuine individual differences in learning ability. Such other factors as the degree of tameness, degree of playfulness, and age differences may have been somewhat effective. The monkeys were all quite tame by the time they were put upon the advanced problems. None of them were particularly playful in the apparatus except subject *A*, which tended to become more so from step to step. This animal did not advance beyond Step VI, and the increasing tendency to play may have prevented him from advancing further. The age factor may be disregarded, since there is no evidence that the older (larger) monkeys learned any more effectively than the younger.

The records show that a given animal was often not very consistent in learning score from step to step. A low score might be made on one step, a high score on the next step, a low score on the following step, etc. Such inconsistency may have been directly related to specific reaction tendencies of the animal, but no direct evidence on this point is available. It may have been partly due, on the other hand, to differences in transfer effect from step to step. A high score on one step, for example, might involve an overlearning of the pattern up to this point which might, in turn, transfer to the next step and thus produce a low score. A low score on a given step might conceivably show less transfer to the next step. This principle of explanation cannot be generally applied, however, since there was by no means a perfect alternation of high

and low scores from step to step for the same animal.

The results relating to limits of learning for the several animals have been summarized in Table 12. The

TABLE 12

SHOWING THE LIMITS OF LEARNING FOR EACH MONKEY

Animal	Step
A	VI
B	XIV
C	XV
D	V
E	XI
F	VIII

range for the group was from step 5 to step 15 inclusive. The median number of steps learned was 9.5 and the average was 9.8. It may be remarked that there was little or no relationship between speed of learning a given step and the limit of learning finally reached. It appears, therefore, that the score number of trials to learn a step and the limits score are measures of different behavioral functions. As suggested by Fjeld (8), the two kinds of score are roughly analogous to speed and power measures in human testing. This interpretation seems to be favored by the work of Riess (18) on rats and guinea-pigs. He found that guinea-pigs were unable to master Step II, while rats succeeded in doing so. However, the guinea-pigs learned Step I with fewer trials than did the rats. Moreover, this view is in harmony with the generally accepted principle that complex tasks are more effective than simple tasks in bringing out individual differences in learning capacity.

No evidence appeared that, after a number of steps had been mastered, the subjects were able to develop a generalized habit of adding another plate to the pattern. If this had occurred at any point it would have meant very low scores on all the later steps. The lowest score, made by subject *A*, was eight trials in passing from Step I to Step II. As shown in Tables 1 and 10 this animal made many high scores in connection with the learning of later steps. The same is true of subject *C* after having learned Step V. There is thus little or no evidence that "insight" or "rational" factors played a part in the mastery of either the basic or advanced problem.

ACTIVITY SCORES

The same kinds of activity scores were secured on the advanced problem as on the basic problem. Scores obtained directly from the original data include number of plates depressed and the number of quadrants traversed. These values will be found in the second and third columns of Table 13. The measures of variability have not been included since the group, small to begin with, was still further reduced from step to step. It is obvious that such measures can have no significance when the group includes only two or three animals. As a matter of fact the individual scores and averages for plates depressed and quadrants traversed appear to show no significant trends from step to step.

The computed scores (quadrants per plate, plates per minute, quadrants per minute) are given in the

TABLE 13

RESULTS OF ADVANCED PROBLEM SHOWING INDIVIDUAL ACTIVITY SCORES AND MEASURES OF CENTRAL TENDENCY FOR NUMBER OF PLATES DEPRESSED, NUMBER OF QUADRANTS TRAVERSED, NUMBER OF QUADRANTS PER PLATE, NUMBER OF PLATES DEPRESSED PER MINUTE, AND NUMBER OF QUADRANTS PER MINUTE

Animals	Plates depressed	Quadrants traversed	Quadrants per plate	Plates per minute	Quadrants per minute
		Step IV			
A	3226	3801	1.18	17.27	20.35
B	307	356	1.16	25.52	29.59
C	4050	5667	1.40	11.99	16.78
D	1859	3896	2.10	4.81	10.09
E	587	805	1.37	18.77	25.74
F	1429	1837	1.29	9.92	12.75
Median	1644.00	2819.00	1.33	14.63	18.57
Average	1909.67	2727.00	1.42	14.71	19.22
		Step V			
A	1130	1238	1.10	23.40	25.63
B	159	160	1.01	29.72	29.91
C	154	182	1.18	26.19	30.95
D	1011	1207	1.19	18.71	22.34
E	371	423	1.14	26.37	30.06
F	380	396	1.04	26.03	27.12
Median	375.50	409.50	1.12	26.11	28.52
Average	534.17	601.00	1.11	25.07	27.67
		Step VI			
A	1802	1905	1.06	26.01	27.50
B	1160	2036	1.76	13.88	24.35
C	731	800	1.09	21.08	23.07
D	Failed				
E	1057	1122	1.06	31.29	33.21
F	1098	1175	1.07	24.36	26.06
Median	1098.00	1175.00	1.07	24.36	26.06
Average	1169.60	1407.60	1.21	23.32	26.84
		Step VII			
A	Failed				
B	614	695	1.13	25.78	29.18
C	906	987	1.09	30.05	32.74
E	774	822	1.06	39.03	41.45
F	2164	2477	1.14	22.71	26.00
Median	840.00	904.50	1.11	27.92	30.96
Average	1114.50	1245.25	1.11	29.39	32.34
		Step VIII			
B	1846	5452	2.95	12.02	35.51
C	560	586	1.05	29.71	31.09
E	1139	1204	1.06	40.18	42.47
F	3739	5384	1.44	10.85	15.63
Median	1492.50	3294.00	1.25	20.87	33.30
Average	1821.00	3156.50	1.63	23.19	31.18

TABLE 13 (*Continued*)

Animals	Plates depressed	Quadrants traversed	Quadrants per plate	Plates per minute	Quadrants per minute
		Step IX			
B	378	1511	4.00	8.85	35.39
C	6946	8076	1.16	30.01	34.89
E	1159	1246	1.08	44.66	48.02
F	Failed				
Median	1159.00	1511.00	1.16	30.01	35.39
Average	2827.67	3611.00	2.08	27.84	39.43
		Step X			
B	1062	1714	1.61	21.92	35.38
C	1504	1543	1.03	43.76	44.89
E	2046	2143	1.05	47.99	50.27
Median	1504.00	1714.00	1.05	43.76	44.89
Average	1537.33	1800.00	1.23	37.89	43.51
		Step XI			
B	301	308	1.02	34.92	35.73
C	778	796	1.02	39.29	40.20
E	4892	5143	1.05	45.30	47.62
Median	778.00	796.00	1.02	39.29	40.20
Average	1990.33	2082.33	1.03	39.84	41.18
		Step XII			
B	2216	2845	1.28	29.67	38.10
C	791	816	1.03	43.58	44.96
E	Failed				
Median	1503.50	1830.50	1.16	36.63	41.53
Average	1503.50	1830.50	1.16	36.63	41.53
		Step XIII			
B	812	825	1.02	42.89	43.58
C	1440	1561	1.08	45.93	49.79
Median	1126.00	1193.00	1.05	44.41	46.69
Average	1126.00	1193.00	1.05	44.41	46.69
		Step XIV			
B	5237	6624	1.26	33.68	42.61
C	1118	1146	1.03	48.97	50.20
Median	3177.50	3885.00	1.15	41.33	46.41
Average	3177.50	3885.00	1.15	41.33	46.41
B	Failed				
		Step XV			
C	1826	2205	1.21	41.27	49.83
		Step XVI			
C	Failed				

last three columns of Table 13. In general, these computed scores seem to indicate degree of adjustment of the animal to the situation. As will be seen the av-

erages for these several scores show no consistent trends from step to step. This might be taken to mean that each step presented to the animal a relatively novel situation. It may be pointed out, moreover, that these several types of scores vary greatly from animal to animal on the same step. Furthermore, an animal might be above the average in activity on one step and below the average of the group on another step. It seems quite impossible, on the basis of the present data, to relate these various types of activity to the learning scores.

An analysis of the errors into various types is summarized in Table 14. In general, the errors here are in the same order as those noted under the basic problem. In addition to these, the error of hesitating on one or another of the plates was checked. This type of error became more and more important as the length of the series was increased. This was due in part to the growing tendency of the animal to stop on the final plate of a series every time it was touched in running off a pattern. For example, on Step VIII which ended with plate 2 hesitation might occur on each of the four times that this plate was depressed. The monkeys often showed a tendency, also, to hesitate on the final plate for the preceding step. Both types of hesitation are combined in one score in the table. A glance at Table 14 will show that the two dominant types of errors throughout the advanced problem were (*a*) reactions to wrong plates and (*b*) hesitation on plates. In a few instances going around clockwise and omissions of

TABLE 14

TYPES OF ERRORS, SHOWING THE DISTRIBUTION OF EACH ON THE SEVERAL STEPS OF THE ADVANCED PROBLEM

Steps Number of animals	IV (6)	V (6)	VI (5)	VII (4)	VIII (4)	IX (3)	X (3)	XI (3)	XII (2)	XIII (2)	XIV (2)	XV (1)
Types of errors												
Omission of plates:												
Plate 1	3.18	4.52	3.39	5.54	2.55	1.92	3.72	11.20	2.53	.91	3.93	.99
Plate 2	.89	2.58	1.00	6.25	4.57	12.29	4.99	6.85	2.85	21.00	6.68	15.54
Plate 3	1.34	.57	1.31	.38	.09	1.00	.85	.46	1.58	1.83	.46	.66
Wrong plates:												
Plate 1	6.47	14.06	20.11	11.90	6.37	18.43	11.92	5.00	4.11	7.99	5.30	36.69
Plate 2	27.13	21.02	22.63	19.77	14.87	16.63	17.08	20.56	15.35	21.00	12.99	13.55
Plate 3	21.43	13.70	5.53	14.83	11.35	10.93	6.09	16.76	12.34	16.21	10.43	21.65
Hesitation on plates:												
Plate 1	.98	10.83	7.50	1.58	.82	11.19	11.41	4.26	1.11	14.38	5.22	1.49
Plate 2	17.26	12.99	14.72	6.90	12.60	1.33	10.99	6.02	23.41	7.99	11.44	4.30
Plate 3	2.59	.79	.62	16.84	2.12	.18	.42	7.59	4.43	.69	1.01	.83
Wrong direction	5.29	10.69	5.77	6.57	4.34	6.47	6.34	15.37	5.38	2.06	5.95	.50
Irregular course	.89	.29	.21	.11	.82	.29	.08	.00	.16	.69	.37	.00
Climbing	2.22	1.29	.90	.54	3.27	.56	.08	.74	.00	1.37	1.10	.33
Entrance door	3.58	1.36	.90	.43	7.28	1.42	4.48	.00	.32	.00	2.74	.00
Inner door	3.56	2.80	3.35	2.55	5.71	7.88	.25	.65	.32	.23	.82	.50
Going around:												
Counterclockwise	.06	.36	.07	.05	.20	.03	.08	.00	.16	.00	.18	.00
Clockwise	.04	.14	6.08	.49	19.95	8.24	12.68	.00	22.63	.00	26.08	.00
Running beyond	1.60	1.08	4.80	4.62	2.52	.92	8.54	4.54	3.32	3.65	5.30	2.97
Miscellaneous	1.49	.93	1.11	.65	.07	.29	.00	.00	.00	.00	.00	.00
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

plates 1 or 2 take first or second place. It would seem, therefore, that the crux of the advanced problem was the selection of the proper ending plate and the elimination of the tendency to hesitate on plates, i.e., reaching certainty as to where and when to stop. The numerous other aspects of the problem, as reflected in the different types of error, appeared to offer relatively little difficulty.

Something can be said regarding the manner in which patterns were developed by the animals. Data on this point were secured in connection with corrected trials, i.e., those in which the pattern was broken at some stage before completion by the intrusion of errors of one sort or another. For example, on Step V the animal might depress plates 1, 2, 3 in order and then run around the incentive compartment before reversing and stepping on plates 2 and 1. These breaks might occur, in individual cases, at any point in the pattern. An analysis of our records showed the same three stages of pattern formation previously noted by Fjeld. These stages, which appeared on each step of the advanced problem, may be summarized as follows: (1) a stereotyped repetition of the pattern of the preceding step marked by hesitation on the final plate, (2) a break-down of this older pattern into units of one or a few plates, (3) a reorganization of these units into the larger pattern which included the additional plate for the new step. This analysis would seem to offer additional evidence against the presence of "insight" in the solution of the advanced problem.

An interesting relationship between the development of the patterns and final failures seemed to exist. Failures on all of the higher steps can be grouped into two classes: those in which the subjects made progress in learning and were nearly always successful in obtaining the food, but never succeeded in reaching the required norm of mastery, as in the cases of *A, C, E,* and *F;* and those in which they first made progress in learning and then failed through disintegration of the habit to the point where they consistently made failure trials, as in the cases of *B* and *D.*

Those monkeys which came under the first class of failures, without exception, did not progress in the step which they failed to master beyond the second stage in pattern development. The animals falling under the second classification showed some evidence of developing both the second and third stages of the pattern, but progress was retarded by the appearance of such antagonistic patterns as starting in the wrong direction and running around the cage without stepping on the plates. Finally, this behavior gave way to the making of failure trials and a subsequent disintegration of the behavior to inactivity.

V

GENERAL COMPARISONS

As noted in the Introduction, this study is a part of a larger project on the limits of learning in mammals which has been under way for some years in the Columbia laboratory. Up to the present, this project includes the work of Shuey (19) on kittens, that of Riess (18) on white rats and guinea-pigs, and that of Fjeld (8) on rhesus monkeys. In the present section it will be desirable to compare our results with those obtained earlier by means of the same method. In addition to these comparisons, a more detailed analysis of the data at hand on cebus and rhesus monkeys will be made.

COMPARISON OF MAMMALS

The limits reached, as indicated by the number of steps learned, by the several species of mammals are given in Table 15. The number of animals of the various groups, expressed in percentages, which learned the several steps is shown in Table 16. The size of the groups, in absolute terms, on the three steps of the basic problem is indicated in Table 17.

TABLE 15

	Range	Median	Average
Guinea-pigs	0- 1	1.0	0.5
Rats	0- 2	1.0	0.9
Kittens	1- 7	3.0	3.6
Rhesus monkeys	2-22	5.0	7.4
Cebus monkeys	5-15	9.5	9.8

TABLE 16

COMPARISON OF TWO TYPES OF MONKEYS, KITTENS, RATS, AND GUINEA-PIGS SHOWING PERCENTAGE OF SUBJECTS WHICH LEARNED THE VARIOUS STEPS

Steps learned	Cebus monkeys	Rhesus monkeys	Kittens	Rats	Guinea-pigs
I	100.0	100.0	100.0	68.6	53.3
II	100.0	100.0	100.0	22.9	
III	100.0	92.9	100.0		
IV	100.0	64.3	20.0*		
V	100.0	50.0	20.0		
VI	83.3	50.0	10.0		
VII	66.7	42.9	10.0		
VIII	66.7	42.9			
IX	50.0	35.7			
X	50.0	28.6			
XI	50.0	28.6			
XII	33.3	28.6			
XIII	33.3	14.3			
XIV	33.3	7.1			
XV	16.7	7.1			
XVI		7.1			
XVII		7.1			
XVIII		7.1			
XIX		7.1			
XX		7.1			
XI		7.1			
XXII		7.1			
XXIII					

*This is the result for Step V which corresponds to Step IV for the monkeys.

As will be seen from Tables 15 and 16, marked differences in ability to form complex patterns of behavior were exhibited by the various animals investigated. The rodents rank lowest in this capacity, the white rats making a better showing than the guinea pigs. Approximately 23 per cent of the rats were able to learn Step II, while all of the guinea-pigs failed on this step. The kittens proved to be markedly superior to either of the rodents. However, only 10 per cent of

TABLE 17

SHOWING MEASURES OF CENTRAL TENDENCY AND VARIABILITY IN TRIALS AND MINUTES FOR TWO TYPES OF MONKEYS, KITTENS, RATS, AND GUINEA-PIGS ON THE VARIOUS STEPS OF THE BASIC PROBLEM

	Median	Average	Range	Average deviation	Standard deviation	Coefficient of variation
Step I (6 C. Monkeys, 17 R. Monkeys, 62 Kittens, 24 Rats, and 16 Guinea-Pigs)						
Trials						
Cebus monkeys	82.00	137.17	42-327	80.83	108.41	79.03
Rhesus monkeys	191.00	162.47	19-310	77.94	94.36	58.08
Kittens	42.14	46.69	9-136	17.80	25.28	54.13
Rats	196.00	221.04	30-453	99.97	125.26	56.67
Guinea-pigs	151.50	185.50	53-407	110.69	176.28	95.03
Minutes						
Cebus monkeys	51.94	58.16	23.07-97.22	28.23	29.61	50.91
Rhesus monkeys	210.88	306.71	17.90-970.87	222.24	279.14	91.01
Kittens	47.00	63.31	8.70-238.00	38.73	51.35	81.10
Rats	100.37	129.39	12-230	77.09	96.59	74.65
Guinea-pigs	74.40	83.28	40-191	39.19	49.10	58.96
Step II (6 C. Monkeys, 17 R. Monkeys, 55 Kittens, 8 Rats)						
Trials						
Cebus monkeys	76.50	77.67	8-135	37.67	45.06	58.01
Rhesus monkeys	60.00	169.65	12-609	139.53	189.25	111.55
Kittens	18.50	22.59	1-70	14.01	17.23	76.27
Rats	365.00	315.25	160-450	93.19	116.77	37.04
Minutes						
Cebus monkeys	17.60	32.91	3.23-124.12	24.84	41.69	126.68
Rhesus monkeys	55.57	70.00	13.65-203.73	42.27	52.12	74.46
Kittens	7.50	12.05	1-71.80	9.37	13.05	108.34
Rats	83.26	81.00	32-139	24.35	31.14	38.44
Step III (6 C. Monkeys, 16 R. Monkeys, 52 Kittens)						
Trials						
Cebus monkeys	81.50	100.83	34-280	50.50	83.15	82.47
Rhesus monkeys	61.00	200.00	1-1640	183.63	397.88	198.94
Kittens	33.75	39.42	1-121	21.91	27.37	69.43
Minutes						
Cebus monkeys	18.32	25.08	9.83-71.02	13.47	21.15	84.33
Rhesus monkeys	15.04	57.01	12-504.07	51.29	120.32	211.05
Kittens	19.20	28.46	2-107.50	20.07	26.45	92.93

the kittens did as well as 83.3 per cent of the cebus monkeys and 50 per cent of the rhesus monkeys. Furthermore, 66.7 per cent of the cebus monkeys, and 42.9 per cent of the rhesus monkeys went beyond Step VII, which corresponds to the upper limit of the kittens. In interpreting these comparisons, it should be borne in mind that the several species of mammals tested were of approximately the same relative age. In every case young animals were employed, and a serious attempt was made to select them with reference to the same point on their respective growth curves.

It may be noted that Step I proved to be difficult for the kittens and for both types of monkeys. The same is true also of Step IV which involved the first reversal of direction. The relative difficulty of the steps beyond this point follows no definite order for either the kittens or the two types of monkeys. Moreover, there is no consistent relationship between the learning scores on the several steps and the limits finally reached by the individual animal. Furthermore, the species showing the highest limit score often required more effort (trials, time) to learn the easier steps than did the species making a lower limit score. For example, the white rats made a higher learning score on Step I than did the guinea-pigs. In general, the learning scores for the monkeys are much higher than those of the kittens. These facts appear to support the view that the scores on the several steps are measures of one thing and the limit scores of another. The latter would seem to be an index of the general level of learning

capacity in the individual and the species. This conclusion would be more certain if corresponding differences were found among these several species on other types of tasks. Perhaps it should be pointed out that the results secured by this method are roughly, at least, in harmony with the opinion of competent animal workers as to the relative intelligence of the species compared.

Comparison of Cebus and Rhesus Monkeys

The main interest in the present problem was to secure data on the cebus monkey that would be strictly comparable to that of Fjeld (8) on the rhesus monkey. In order to make such a comparison possible the testing of both types was standardized in the following particulars: (*a*) the animals were kept in the same laboratory, given the same care, and tamed in the same manner; (*b*) the tests were conducted in the same dark room, using identical apparatus and technique; (*c*) the apparatus itself was one that had been proved to be of value in the determination of levels of learning ability in different animals; and (*d*) the kinds of scores taken and the methods of computing them were the same in both studies. Our experiment was thus so planned as to bring out differences in level of learning capacity between the cebus and the rhesus monkey, if such were present.

When the results are taken at their face value it is by no means easy to say which type of monkey, if either, is the superior. As indicated in Table 16, the best

rhesus monkey stands considerably higher than the best cebus monkey. On the other hand the worst rhesus monkey stands considerably lower than the lowest cebus monkey. In brief, the range for the cebus group lies wholly within that of the rhesus group. The median number of steps learned by the cebus monkey was 9.5 and the average was 9.8. The corresponding values for the rhesus group were 5.0 and 7.4. On the average, then, the cebus monkey did considerably better than the rhesus monkey.

The latter group included seven males and seven females. According to Fjeld, the rhesus males proved to be superior to the females in terms of limit scores. In fact, only 14.29 per cent of the females reached or exceeded the median of the males. It would be fairer, perhaps, to compare the scores of our six male cebus monkeys with those of her seven male rhesus monkeys. The medians for the two groups are almost the same, *being 9.5 for the cebus group and 9.0 for the rhesus* males. The averages are almost identical, 9.8 being the value for the cebus group and 9.6 for the rhesus group. The range of the cebus group is 5 to 15 steps, while that of the rhesus males is 3 to 22 steps. The range of the latter group would be 3 to 12, however, if the record of the best animal in the group was omitted. This animal represents an extreme variation, since it went 10 steps beyond any other male and $1\frac{1}{3}$ steps beyond any female of the rhesus group. An extreme case of this sort would be more likely to appear in a group of 14 monkeys than in a group of six monkeys. Per-

haps, the range for the male rhesus group is more representative with this extreme case omitted than with it included.

In view of these facts, the most logical interpretation would seem to be that no clear difference had been shown to exist between the two types of monkeys in terms of limits of learning. This is not to deny that a genuine difference may be present, but merely to insist that the data at hand do not demonstrate such. If the number of animals were increased by additional work on both types, a real difference in the level of learning capacity might be found. It may be true, for example, that the rhesus monkeys are less homogeneous in this and related functions than the cebus. This would mean a wider spread of limit scores, and the range might well include the range of cebus monkeys. On the other hand, an extreme case like that found in the rhesus group might make its appearance among the cebus monkeys, if more individuals were tested. The fact that the lower limit of the range for the cebus group was two steps higher than that for the rhesus group seems to indicate that the simpler stages of the problem are somewhat easier for the cebus monkeys. This conclusion is corroborated by the fact that the average number of trials to learn on each of the first five steps was approximately twice as great for the rhesus as for the cebus monkey. Beyond Step V the differences are not very great and the averages from step to step do not consistently favor either group.

It is obvious that the groups are so small that a sta-

tistical treatment of reliability is hardly justified. Nevertheless, it may be of some value to give the results of such an analysis for what they may be worth. The reliability of the differences between the averages in limit scores for the several groups is shown in Table 18. As will be seen, there is only one chance in 100

TABLE 18

SHOWING RELIABILITY OF THE DIFFERENCE IN THE AVERAGE LIMIT SCORE FOR THE VARIOUS GROUPS

Groups	Standard deviation of the difference	The difference	Diff./S.D. of difference	Chances in 100 of a true difference
Rhesus M. and F. group with rhesus males	8.24	2.2	.27	10
Rhesus M. and F. group with cebus males	6.66	2.4	.36	14
Rhesus males with cebus males	7.23	0.2	.03	1

that a very slight difference between the rhesus males and the cebus males is a true difference. Since the reliability of the difference is so extremely low, these two groups of males might well be combined into a single group of 13 males. That is, they may be regarded as if they belonged to the same species, insofar as the limit score is concerned. As indicated in the table there are only 14 chances in 100 that the difference between the rhesus group (males and females) and the cebus group is a true difference. It would appear, therefore, that the species distinction may also be ignored when the male and female rhesus are taken together. As a matter of fact the measures of reliability

haps, the range for the male rhesus group is more representative with this extreme case omitted than with it included.

In view of these facts, the most logical interpretation would seem to be that no clear difference had been shown to exist between the two types of monkeys in terms of limits of learning. This is not to deny that a genuine difference may be present, but merely to insist that the data at hand do not demonstrate such. If the number of animals were increased by additional work on both types, a real difference in the level of learning capacity might be found. It may be true, for example, that the rhesus monkeys are less homogeneous in this and related functions than the cebus. This would mean a wider spread of limit scores, and the range might well include the range of cebus monkeys. On the other hand, an extreme case like that found in the rhesus group might make its appearance among the cebus monkeys, if more individuals were tested. The fact that the lower limit of the range for the cebus group was two steps higher than that for the rhesus group seems to indicate that the simpler stages of the problem are somewhat easier for the cebus monkeys. This conclusion is corroborated by the fact that the average number of trials to learn on each of the first five steps was approximately twice as great for the rhesus as for the cebus monkey. Beyond Step V the differences are not very great and the averages from step to step do not consistently favor either group.

It is obvious that the groups are so small that a sta-

tistical treatment of reliability is hardly justified. Nevertheless, it may be of some value to give the results of such an analysis for what they may be worth. The reliability of the differences between the averages in limit scores for the several groups is shown in Table 18. As will be seen, there is only one chance in 100

TABLE 18

SHOWING RELIABILITY OF THE DIFFERENCE IN THE AVERAGE LIMIT SCORE FOR THE VARIOUS GROUPS

Groups	Standard deviation of the difference	The difference	Diff./S.D. of difference	Chances in 100 of a true difference
Rhesus M. and F. group with rhesus males	8.24	2.2	.27	10
Rhesus M. and F. group with cebus males	6.66	2.4	.36	14
Rhesus males with cebus males	7.23	0.2	.03	1

that a very slight difference between the rhesus males and the cebus males is a true difference. Since the reliability of the difference is so extremely low, these two groups of males might well be combined into a single group of 13 males. That is, they may be regarded as if they belonged to the same species, insofar as the limit score is concerned. As indicated in the table there are only 14 chances in 100 that the difference between the rhesus group (males and females) and the cebus group is a true difference. It would appear, therefore, that the species distinction may also be ignored when the male and female rhesus are taken together. As a matter of fact the measures of reliability

are about the same between the total rhesus group and the cebus males as between the total rhesus group and the rhesus males. It is evident, therefore, that in so far as such statistical treatment has any meaning for such small groups, the data show that there is no reliable difference between the limit scores of the cebus and the rhesus monkey. In fact, our six cebus monkeys and the 14 rhesus monkeys of Fjeld might well be thrown together into a single group. It is interesting to note, that when this is done gaps in the distribution curve of the steps are filled in at various points.

After observing a number of rhesus monkeys work at the task, and after testing the cebus monkeys, it appeared to the experimenter that there were fairly definite differences in mode of attack. Apparently the reactions of the cebus were less nervous and somewhat more deliberate in approaching the plates. They seemed to be more docile in temperament and, as has been shown, they learned each of the first five steps in approximately half the number of trials taken by the rhesus monkey.

In spite of these observations, very few differences between the two types are revealed by the activity scores when these are computed on a per trial basis. The activity scores for the cebus monkeys are given in Tables 6 and 13, and those for the rhesus monkeys in Tables 10, 24, and 26 of Fjeld's monograph (8). The ratio of plates depressed and quadrants traversed, while not altogether consistent, is roughly the same for both types. Since quadrants-traversed is an index of general activ-

ity, it appear that the cebus monkeys showed as much exploratory behavior in the apparatus as did the rhesus monkeys. The quadrant per plate ratio was practically the same for both types throughout. The number of plates touched per minute was much greater for the cebus monkeys on the first two steps of the basic problem, but the difference between the two groups was slight thereafter. The same relationship held also for the quadrant per minute score. These two trends indicate that the cebus monkeys moved about the apparatus at a faster rate and depressed more plates per minute on the first two steps. On the whole, however, it would seem that the two types of monkeys were remarkably similar in the forms of behavior represented by the various kinds of activity scores.

An analysis of the types of errors for the cebus will be found in Tables 9 and 13, and for the rhesus in Tables 16 and 28 of Fjeld's monograph (8). The values of these tables are the percentages of the several types of errors when the total error score for the group is taken as 100 per cent. They represent, therefore, the relative frequency of the several types of errors in the learning process as a whole. As previously noted, there were three types of errors made in connection with the plates: omissions, wrong plates, and hesitation on plates. An analysis of the results shows that the cebus monkeys omitted plates much less often than the rhesus monkeys. On the other hand, they made many more wrong-plate errors. Hesitation on plates, which was checked only on the advanced problem, occurred much

less frequently in the case of the cebus monkeys. It seems, therefore, that the cebus monkeys developed a more *definite reaction to* plates than the rhesus monkeys. This is indicated both by the low omission score and by the high wrong-plate score. This definite reaction to plates, as well as the less hesitation on plates, may be explained in part by the more calm and deliberate activity of the cebus monkeys in the apparatus.

Of the remaining types of errors, only a few seem to reveal significant differences between the cebus and rhesus monkeys. The activity of zigzagging about in the apparatus (irregular course) shows a percentage only one-fourth as high for the cebus as for the rhesus. Climbing about the cage also occurred much less frequently in the cebus monkeys. The same is true, as well, of the reaction of sitting at the entrance door. These differences in manner of working are probably related to differences in the characteristic temperaments of the two species. On the whole it would seem that the nervous temperament of the rhesus led to the types of errors associated with exploring the apparatus. On the other hand, the more deliberate attitude of the cebus manifested itself in fewer errors of this kind, with a relative increase of wrong-plate errors. These differences in manner of working at the problem, while interesting in themselves, did not seem to influence greatly the limit score. This may be explained, in part at least, by the fact that both species exhibited the same general method of pattern formation in the advanced problem. This development as already noted led

through the three following stages in connection with each step: (1) repetition of the old pattern, (2) the breaking down of this pattern, and (3) the integration of the new and more complex pattern. It seems likely that the mode of attack was so dominated by this general process of pattern formation that the relative frequency of specific types of errors became a matter of small importance.

The comparisons here made offer no support to the opinion that the cebus monkey is less intelligent than the rhesus. In fact the results obtained, insofar as they go, may be regarded as so much evidence in favor of the view that the level of monkey intelligence is essentially the same in the New World and Old World groups. A definite conclusion to this effect cannot be drawn, however, because only one species from each of the Cebidæ (New World) and Cercopithecidæ (Old World) groups has been tested. There are eight other known varieties of the Cebidæ and five other varieties of the Cercopithecidæ. It seems likely that the cebus and the rhesus may be taken as fairly representative of the higher types of monkeys in their respective families. In any case a beginning has been made toward the direct comparison of the New World and the Old World monkeys as to general level of intelligence. It is recognized, of course, that the apparatus and method here employed do not yield a complex index of intelligence in any animal. The comparative results secured on other kinds of problems would need to be taken into account in forming a final conclusion, even if the latter should be restricted to cebus and rhesus monkeys.

VI
SUMMARY

Six young cebus males were tested on the Jenkins Problem Apparatus, which provides a series of tasks of increasing complexity. Each animal was transferred from one step to the next, as soon as it reached the norm of nine perfect trials out of ten, until it finally failed to learn a given step within 1000 trials. The last step mastered was taken as the limit of learning for the animal. The limit scores, the learning scores (trials, errors, minutes) for each step, and a number of activity scores were computed for the group, and compared with similar data secured by Fjeld (8) on rhesus monkeys. The following conclusions may be drawn from these results:

1. The limits of learning for the six male cebus monkeys range from 5 to 15 steps, with a median of 9.5 and an average of 9.8. The limits of learning for seven male rhesus monkeys, as given by Fjeld, range from 3 to 22 steps, with a median of 9.0 and an average of 9.6; the corresponding values for her combined group (7 males and 7 females) are 2 to 22, 5.0, and 7.43. A comparison of these data led to the tentative conclusion that there is no significant difference between cebus and rhesus monkeys in the level of capacity to form habit patterns of this type. The monkeys are markedly superior, however, to such mammals as the kittens (Shuey: range, 3 to 7; average, 3.6), the white rat (Riess: range, 0 to 2; average, 0.9), and guinea-pigs (Riess: range, 0 to 1; average, 0.5).

2. Minor differences in the behavior of the cebus and rhesus monkey were revealed by an analysis of the various kinds of activity scores and the several types of errors recorded. These differences appear to be due, in the main, to temperamental factors which distinguish the two species. The general method of pattern formation, as the problem becomes complicated, is the same for both species. This process involves the following stages: (1) repetition of the old pattern, (2) the breaking down of this pattern, and (3) the integration of the new and more complex pattern. No evidence was found for the development of a generalized solution to the problem such as that involving "the addition of another plate" from stage to stage.

3. In general, the average number of trials required to learn was much less for the cebus than for the rhesus monkey, the ratio on the first five steps being approximately one to two. In spite of this fact, the average limits for the two species were roughly the same, as indicated above. As in the work of Fjeld, Shuey, and Riess, no significant relationship was found between number of trials required to learn the several steps and the final limit score of the individual. The former appears to measure speed of learning while the latter indicates general level of learning ability.

4. Individual differences were marked both in learning scores and in activity scores for a given step. The performance of the individual, in terms of these indices, also varied greatly from step to step. Neither the individual scores nor the group scores show a consistent tendency to decrease from step to step.

5. The various steps of the problem presented in this apparatus were found not to progress by equal differences in difficulty. Step IV proved to be especially difficult for most of the monkeys, probably because it involved the first reversal of direction in pattern formation. No single step was found to be consistently easy or difficult for all of the animals.

References

1. Bierens de Haan, J. A. Über das Suchen nach unsichtbaren Futter bei Affen und Halbaffen. Zugleich ein Beitrag zu der Frage nach dem konkreten Verstandnis dieser Tiere. *Zsch. f. vergl. Physiol.*, 1930, 11, 630-655.
2. ———. Werkzeuggebrauch und Werkzeugherstellung bei einem niederen Affen (*Cebus hypoleucus Humb.*) *Zsch. f. vergl. Physiol.*, 1931, 13, 639-695.
3. Buytendyk, F. J. J. Considerations de psychologie comparée à propos d'expériences faites avec le singe cercopithecus. *Arch. Neérl. de physiol.*, 1921, 5, 42-88.
4. Carpenter, C. R. Field study of behavior and social relations of howling monkeys. *Comp. Psychol. Monog.*, 1934, 10. Pp. 168.
5. Drescher, K., & Trendelenburg, W. Weiteren Beitrag zur Intelligenzprufung in Affen (einschlusslich Anthropoiden. *Zsch. f. vergl. Physiol.*, 1927, 5, 613-642.
6. Gellermann, L. W. The double alternation problem: I. The behavior of monkeys in a double alternation temporal maze. *J. Genet. Psychol.*, 1931, 39, 50-72.
7. ———. The double alternation problem. III. *J. Genet. Psychol.*, 1931, 39, 359-392.
8. Fjeld, H. A. The limits of learning ability in Rhesus monkeys. *Genet. Psychol. Monog.*, 1934, 15, 369-537.
9. Hartman, C. G., & Straus, W. L., J. [Ed.] The anatomy of the rhesus monkey. Baltimore: Williams & Wilkins, 1933. Pp. 383.
10. Hobhouse, L. T. Mind in evolution. (3rd. ed.) London: Macmillan, 1926. Pp. 483.
11. Jenkins, T. N. A standard problem box of multiple complexity for use in comparative studies. *J. Comp. Psychol.*, 1927, 7, 129-144.
12. Klüver, H. Behavior mechanisms in monkeys. Chicago: Univ. Chicago Press, 1932. Pp. 387.
13. Leche, W. Der Mensch. Jena: Fischer, 1911. Pp. vii+390.
14. Lydekker, R. The royal natural history. Vol. 1. London: Warne, 1893-6. Pp. xvi+583.

15. McDougall, K. D., & McDougall, W. Insight and foresight in various animals—monkey, raccoon, rat, and wasp. *J. Comp. Psychol.*, 1931, 11, 237-274.

16. Nellmann, H., & Trendelenburg, W. Ein Beitrag zur Intelligenzprufung niederer Affen. *Zsch. f. vergl. Physiol.*, 1926, 4, 142-200.

17. Révész, G. Expériences sur la memoire topographique et sur la decouverte d'un systeme chez des infants et des singes inferieurs. *Arch. de psychol.*, 1923, 18, 323-342.

18. Riess, B. F. The limits of learning ability in the white rat and the guinea pig. *Genet. Psychol. Monog.*, 1934, 15, 303-368.

19. Shuey, A. M. The limits of learning ability in kittens. *Genet. Psychol. Monog.*, 1931, 10, 287-378.

20. Thorndike, E. L. Animal intelligence, an experimental study of the associative processes in animals. *Psychol. Rev., Monog. Suppl.*, 1898, 2. Pp. 109.

21. ———. The mental life of the monkeys. *Psychol. Rev., Monog. Suppl.*, 1901, 3. Pp. 57.

22. Tilney, F. The brain from ape to man. Vol. I. New York: Hoeber, 1928. Pp. xxvii+473.

23. Warden, C. J. Animal motivation: experimental studies on the albino rat. New York: Columbia Univ. Press, 1931. Pp. 205.

24. Warden, C. J., Jenkins, T. N., & Warner, L. H. Comparative psychology. New York: Ronald, 1934. Pp. x+581.

25. Woodrow, H. Discrimination by the monkey of temporal sequences of varying number of stimuli. *J. Comp. Psychol.*, 1929, 9, 123-157.

26. Yerkes, R. M. The mental life of monkeys and apes: a study of ideational behavior. *Behav. Monog.*, 1916, 3, No. 12. Pp. 145.

27. Zuckerman, S. Functional affinities of man, monkeys, and apes. New York: Harcourt, Brace. 1933. Pp. 203.

LES LIMITES DE LA CAPACITÉ D'APPRENDRE CHEZ LES SINGES CÉBUS

(Réumé)

On a testé six jeunes cébus mâles sur l'Appareil de Problèmes de Jenkins, qui fournit une série de tâches d'une complexité croissante, pour déterminer leurs limites d'apprentissage. Le problème se compose de déprimer une plaque ou plus dans le plancher du boîte à problèmes. On a commencé par une seule plaque et a augmenté peu à peu la complexité de la réponse en ajoutant des plaques, une chaque fois, jusqu'à ce que l'on ait étendu la série hors de la capacité d'apprendre de l'animal. Aussitôt qu'un animal eut atteint le critère de neuf épreuves parfaites sur dix, on l'a avancé à la prochaine étape, c'est-à-dire, on a ajouté encore une plaque à la série. On l'a fait subir aux animaux pendant la première partie de l'après-midi, et on leur a fait subir dix épreuves par jour en séries de cinq. Le stimulant employé a été un petit raisin sec et un morceau de pomme également petit. On a limité la durée de chaque épreuve à trois minutes. Quand un animal n'a pas réussi à atteindre la norme de maîtrise (neuf épreuves parfaites sur dix) sur une étape donnée en 1000 épreuves, on a cessé l'entraînement. On a devisé l'expérience, pour la rendre plus commode, en deux parties: (*a*) le problème fondamental et (*b*) le problème avancé. La tâche sur le problème fondamental consistait en trois étapes. Dans l'Etape I, les singes ont été entraînés à toucher la plaque 1; dans l'Etape II, les plaques 1 et 2 en ordre, et dans l'Etape III, les plaques 1, 2 et 3 en ordre. Dans le problème avancé, qui n'a été qu'une continuation de la forme hors de la forme de la forme de trois étapes, on a employé le choc comme suggestion dans l'apprentissage des diverses étapes. Le problème avancé a nécessité un renversement de la direction de la part de l'animal, sur les étapes 4, 6, 8, 10, 12, et 14. On a divisé les données obtenues en deux classes, pour les rendre plus commodes, les données de l'apprentissage (épreuves, erreurs, et résultats de temps) et les données de l'activité (plaques déprimées, quadrants traversés, nombre de quadrants pour chaque plaque, nombre de plaques pour chaque minute, nombre de quadrants pour chaque minute et d'autres données de l'activité). Les limites, en termes d'étapes, atteintes par chaque animal sont les suivantes: V, VI, VIII, XI, XIV, XV. La portée des étapes est de 5 à 15, avec une moyenne de 9,8. L'intérêt principal du problème actuel a été d'obtenir des données sur le singe cébus desquelles seraient absolument comparables à celles de Fjeld (8) sur le singe rhésus. Une comparaison de ces données a mené à la conclusion provisoire qu'il n'y a aucune différence signifiante entre les singes rhésus et cébus dans le niveau de capacité de former des formes d'habitude de ce type. Les singes sont cependant très supérieurs à tels mammifères que les chatons, les rats blancs et les cochons d'Inde. Les différences qui semblent distinguer les singes cébus des rhésus sont dues, paraît-il, principalement aux facteurs de tempérament qui distinguent les deux espèces. La méthode générale de la formation des habitudes, comme le problème devient plus compliqué, est la même pour les deux espèces. Ce processus comprend les étapes suivantes: (1) la répétition de l'ancienne forme, (2) la destruction de cette forme, et (3) l'intégration de la forme nouvelle et plus complexe. On n'a trouvé aucune évidence pour le développement d'une solution généralisée du problème telle que celle où il s'agit de "l'addition d'une autre plaque" d'étape en étape.

Koch

DIE GRENZEN DER LERNFÄHIGKEIT BEI CEBUSAFFEN

(*Referat*)

Sechs junge Cebusmännchen wurden im Jenkins-Problemapparat untersucht, der eine Reihe Aufgaben von zunehmender Kompliziertheit ermöglicht, um die Grenzen ihrer Lernfähigkeit zu bestimmen. Das Problem besteht daraus, dass ein Tier eine oder mehere Platten auf dem Boden des Problemkastens drückt. Es wurde mit einer einzelnen Platte angefangen, und die Kompliziertheit der Reaktion wurde durch die Hinzusetzung von Platten eine nach der anderen allmählich vermehrt, bis die Reihe über die Lernfähigkeit des Tieres hinausging. Sobald das Tier das Kriterion von neun fehlerlosen Versuchen aus zehn erreichte, ging es eine Stufe höher, d.h. eine andere Platte wurde hinzugesetzt. Die Tiere wurden während des Vormittags geprüft und erhielten zehn Proben pro Tag in einer Reihe von fünf. Die Belohnung war eine kleine Rosine und ein gleich kleines Stück Apfel. Die Länge jeder Probe wurde auf drei Minuten gesetzt. Wenn es einem Tier misslang, die Norm der Bemeisterung (neun fehlerlose Proben aus zehn) auf einer gegebenen Stufe innerhalb 1000 Versuche zu erreichen, wurde der Versuch aufgegeben. Das Experiment wurde um der Bequemlichkeit willen in zwei Teile eingeteilt: (*a*) das Grundproblem und (*b*) das fortgeschrittene Problem. Die Aufgabe beim Grundproblem bestand aus drei Stufen. Bei Stufe I wurden die Affen trainiert, Platte 1 zu berühren; bei Stufe II, Platten 1 und 2 der Reihe nach, und bei Stufe III, Platten 1, 2 und 3 der Reihe nach. Bei dem fortgeschrittenen Problem, das bloss eine Fortsetzung des Gebildes jenseits des dreistufigen Musters war, wurde Schock als Ansporn zum Lernen der verschiedenen Stufen gebraucht. Das fortgeschrittene Problem erforderte eine Umkehrung der Richtung seitens des Tieres bei Stufen 4, 6, 8, 10, 12, und 14. Die erhaltenen Werte wurden in zwei Klassen um der Bequemlichkeit willen eingeteilt: Lerndaten (Versuche, Irrtümer, und Zeitwerte) und Tätigkeitsdaten (niedergedrückte Platten, durchlaufene Quadranten, Anzahl der Quadranten pro Platte, Anzahl der Platten pro Minute, Anzahl der Quadranten pro Minute und andere Tätigkeitsdaten). Die Grenzen in Ausdrücken der Stufen, die die Tiere erreichten, waren folgende: V, VI, VIII, XI, XIV, XV. Die Rangstufe in Stufen ist von 5 bis 15 mit einem Durchschnitt von 9,8. Das Hauptinteresse bei dem vorliegenden Problem war die Erhaltung von Daten über den Cebusaffen, die mit denen von Fjeld (8) streng vergleichbar wären. Ein Vergleich dieser Daten führte zu dem probenden Schluss, dass es keinen bedeutsamen Unterschied zwischen Cebus- und Rhesusaffen auf dem Niveau der Fähigkeit gibt, Gewohnheitsgebilde dieses Typus zu bilden. Die Tiere aber sind solchen Säugetieren, wie Kätzchen, weissen Ratten und Meerschweinchen, merklich überlegend. Die Unterschiede, die sich zur Unterscheidung des Cebus- vom Rhesusaffen zeigen, erscheinen zum grossen Teil als Folge der Temperamentsfaktoren, welche die beiden Arten unterscheiden. Die allgemeine Methode zur Gebildegestaltung, wie das Problem verwickelt wird, ist dieselbe bei den beiden Arten. Dieser Prozess besteht aus den folgenden Stadien: (1) Wiederholung des alten Gebildes, (2) die Aufbrechung dieses Gebildes, und (3) die Integration des neuen und verwickelteren Gebildes. Keine Evidenz zur Entwicklung einer verallgemeinerten Lösung des Problems zeigte sich, wie diejenige, die aus der "Hinzusetzung noch einer Platte" von Stufe zu Stufe bestand.

KOCH

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

August, 1935
Volume XVII, No. 4

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

August, 1935
Volume XVII, No. 4

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

NATURE-NURTURE AND INTELLIGENCE

From the Institute of Child Welfare, University of Minnesota

By

ALICE M. LEAHY

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

Acknowledgments

The writer is greatly indebted to the Minnesota Board of Control and to its Children's Bureau for making accessible their files, which were necessary in the location of our subjects. She is also grateful to the Minneapolis and St. Paul schools, whose cooperation facilitated the examination of the children composing the control group.

To the graduate assistants of the Institute of Child Welfare, University of Minnesota, who carried out many clerical tasks in connection with the study, and to Miss Blanche Anderson, statistical clerk, she expresses sincere thanks. To Miss Winona Morgan, whose able assistance was unfailing, the deepest appreciation is expressed. To Dr. Florence L. Goodenough, whose thoughtful suggestions were most valuable, she is also indebted. To Dr. John E. Anderson, whose constructive criticism and guidance persisted throughout the study, the highest praise is extended.

Finally, the writer wishes to thank the hundreds of parents and children whose gracious cooperation made the study a reality.

ALICE M. LEAHY

UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MINNESOTA

CONTENTS

I

INTRODUCTION

The Problem

Variation in human intelligence is universally recognized. But experimentation to discover the causes which affect this variation has moved slowly. The reasons are obvious. First, conditions which permit the control of either heredity or environment are difficult to secure, and secondly, our tools for measurement are limited and crude. Although identical twins provide an absolute control of heredity, their separate location in diverse environments is rare. Experimentation involving the control of environment, on the other hand, is not entirely possible. Measures are available for only certain of its features. For its dynamic attributes we have no measures. Hence, what may appear to be similar environments are only approximately identical. However, the individual mental examination has been demonstrated to be fairly reliable of what may be called test intelligence.

The present investigation approaches the problem by a comparison of two groups of children living in approximately identical environments. In one group, the children are unrelated by blood or marriage to the persons shaping the environment. They are adopted children. In the other group, the children are the offspring of the persons who have shaped the environment. Both heredity and environment are operative in the latter group, while in the former, only environment.

Resemblance as expressed by means of the correlation between attributes in the home and test intelligence of child will constitute one type of analysis. Presumably the magnitude of the correlation between adopted children and their foster homes is a function of environment. In the case of parents and true children, it is a function of heredity and environment combined.

A comparative analysis of mean intelligence with cultural levels will constitute a second type of analysis. Since, as will subsequently be shown, the mean intelligence of the two groups of children is almost identical, marked contrasts in intelligence under constant environmental conditions would place the burden of causation on heredity.

If random placement of adopted children exists in each social stratum then variation from the mean intelligence of the group may be assigned to environmental diversity. Since the homes in both groups of children represent an equal spread on the cultural scale from high to low, we may assume that the nurture factor is equally potent in the determination of final variance in IQ. If the genetic variation in intelligence of each group of children corresponds to the variation measured by our test of intelligence, then the contribution of environment toward final variability in IQ of adopteds would be definitely increased, since they are somewhat less variable than the children reared by their own parents.

It should be emphasized that whatever trends and conclusions can be found in this study are valid only

for populations as homogeneous in racial extraction, social standards, and educational opportunities as that from which our subjects are taken. The distribution of homes of the children in this investigation are probably somewhat skewed toward a superior level. Adoptive homes of even the lowest occupational and economic levels are undoubtedly superior in respect to other traits, since society's control and imposition of standards on this type of home is much greater than on the ordinary home. The educational requirement adhered to in matching our adoptive homes with homes in the general population would tend to raise the environmental and genetic level of the homes of the latter group. This would be particularly true in the lowest occupational groups. In the main, the homes were as variable in essential features as homes of an American urban white population. Clearly they were not as variable as if the homes of southern negroes and poor mountain whites had been included. In consequence, home environment cannot be expected to have as large a proportional effect upon the mental differences of the children studied as though they were being reared in unselected families.

However, attention should also be drawn to the fact that the distribution of inheritable mental capacity of the children in this investigation was probably skewed toward a superior level. No children of the idiot or imbecile grade are included. The true parents of the adopted children were somewhat superior in cultural status to parents of dependent children in general. Hence, heredity cannot be expected to contribute as

large a proportional influence to the mental differences of the children as though a greater variation in genetic intelligence was included. Since environment was equally variable in both the experimental and control populations, and since our sample of parents and true offspring (control population) consistently yielded coefficients of resemblance of .50, it is fair to assume that no serious understatement of the general influence of environment exists in our experimental data.

II

HISTORICAL REVIEW

Earlier Studies

Previous to the present investigation, adopted or foster children have twice been used in major studies of nature-nurture and intelligence. Burks (1) of Stanford and Freeman (2) of Chicago published their notable investigations in 1928. In fact, the discrepancies between the results of their investigations stimulated the present research.

Burks undertook a comparative analysis of foster parent-foster child resemblance and true parent-true child resemblance. The children of her two populations were matched for sex, age, and occupational status of father. For the foster group she found consistently low correlations, averaging about .20 between the intelligence of foster children and various attributes measured in their foster homes; while for her control group of true parents and children the coefficients for the same factors were consistent with those established for filial resemblance in physical traits, i.e., about .50. The simultaneous operation of heredity and environment is posited as the explanation for the latter, while the operation of environment alone is offered as the explanation of the coefficients between foster parents and children.

The logical force of the last inference hinges on the freedom of her foster group from the probability of selective mental resemblance in placement. If mental resemblance between foster parents and children ob-

tains, then the relationship exhibited may be due to it rather than to any effect of environment.

Burks circumscribed her foster group by limiting it to children who had been adopted previous to the age of 12 months. Thus the possibility of precise judgments of mental ability on the basis of overt behavior was definitely minimized. Further, she restricted her population to the white race and excluded south European and Jewish children. From her analysis of the social data available on family background, she concluded that the possibility of selective mental resemblance between foster parent and child on the basis of cultural status did not exist.

In contrast to the low coefficients of correlation between test intelligence of child and the foster home found by Burks, Freeman secured coefficients that ranged from .32 to .52 when certain subclassifications were used, and .48 for his entire population. From this evidence he concludes that environment is capable of exercising an influence on mental ability commensurate with that established for true parent and child in which both heredity and environment are operative.

Although Freeman's method was essentially the same as Burks's, his experimental group was not matched with a control group. His foster population also differed considerably in composition. It was practically double in size (401); age at placement ranged from 6 months to 17 years, with the mean at 4 years, 2 months; 8.47 per cent were negro children; four years' residence in the foster home in which the child was located permitted admission into the experimental group.

The possibility of superior foster parents selecting initially bright children was considered a small factor in the relationship, by the author. Arguments were advanced in support of a negligible selective influence because: (1) in 82 per cent of the cases no mental test had been given before placement; (2) adequate family histories of children were not available by which their mentalities might have been estimated; (3) the average age (4 years, 2 months) at which these children were placed in foster homes precluded dependable estimates of mental ability from observable behavior; (4) when Negro children were omitted, the coefficient of correlation between IQ and foster home rating remained practically the same; (5) when 156 children, for whom it would have been least possible to estimate their intelligence since they had been placed under the age of 2 years, were considered independently, a coefficient of $.52 \pm .04$ was secured between test intelligence and foster home rating; (6) when 59 children for whom *no histories* were available were considered separately from the entire group, a correlation of $.50 \pm .06$ between test intelligence and foster home rating was obtained.

In considering Freeman's explanation for the absence of selective placement in his data, one would promptly agree that estimates of mental ability in which no objective test data enter are distinctly limited. At age 4 years, however, the possibility of estimates of mentality from observable behavior is not remote, especially with children who have been under observation for any length of time. According to the

author's tables, the legitimate children, who constituted 65 per cent of the total foster group, were known to the placement agency and cared for in its receiving home or in a temporary boarding home before placement in permanent foster homes for a period of 11 months, on the average. The illegitimate children, who constituted 35 per cent of the group, were known and under care for a period of 3 months, on the average. The mean age at permanent placement for the legitimate adopted children was 5 years, 8 months; for the illegitimate children, 1 year, 7 months, and as previously stated the average age at placement for the entire group was 4 years, 2 months.

Whether the family histories in the Freeman population were sufficiently complete to forecast the potentialities of the children can be seriously doubted. Nevertheless, the amount of information reported caused the author to conclude that "the foster children came from inferior homes and had a heredity which was decidedly poor." It is possible that these same data supplemented by other unrecorded information influenced the placement agency in its choice of foster homes. In so far as this occurred, the reported coefficient of correlation between test intelligence of children and foster home rating is weighted by selective placement. If all sources of selective placement were removed in the Freeman population, then the correlation $.48 \pm .03$, found for test intelligence and foster home rating when Negro children were omitted, would stand as substantial evidence in support of the influence of environment on the mental development of children.

It is apparent from our analysis of the foregoing studies that a knowledge of selective placement is essential to the interpretation of nature-nurture findings in which foster or adopted children are the subjects. A mathematical expression of resemblance between the intelligence of adopted children and their homes may be a function of either selective placement or environment, or of selective placement and environment combined in unknown proportions. Conclusions cannot be made for either nature or nurture, unless it can be demonstrated that the experimental population is free from the effects of selective mental resemblance.

Present Study

In the formulation of the present investigation the problem of mental resemblance between foster parent and child resulting from selective placement received first consideration. Could it be controlled? Obviously if reliable preplacement tests were available the part that selective placement plays in the choice of a home for a child or a group of children could be definitely determined. Moreover, the actual measurement of gains and losses in performance on standard tests of mental ability accordingly as a child is exposed to extremes in environmental stimulation would then be possible. However, such an experiment would assume not only reliable preplacement tests of the child but equally reliable measures of environment. Neither are available. The files of child placement agencies reveal that only a small number of adopted children are

tested in advance of placement. Out of a total of 2449 children adopted in Minnesota between the years 1918-1928, only 98 had been given mental tests. Most of these were tested subsequent to placement. In fact, failure to adjust in the new home generally prompted the examination.

Despite the absence of prognostic tests of mental ability it is highly probable that intellectual promise judged from overt behavior and family history enters into a social agency's recommendations for adoption. The questionnaire replies of 22 child placement workers revealed that 18 regarded *probable intelligence* as of *very great* significance in their judgments of the fitness of a child for an adoptive home, two regarded it of *great* importance, and two of *slight* significance on a scale of five descriptive levels, namely: very slight significance, slight moderate, great, very great significance. Certainly such judgments would fall farthest from the mark in the case of infants. More accurate prediction of future development on the basis of overt behavior could conceivably be made for older children. Indeed very careful placement might result in striking similarity in the intellectual level of foster parents and their adopted children.

The problems that confront the investigator in a research population of untested children placed at older ages are two. First, he must know what elements of history and behavior contributed to the judgment of the child's mental ability and, secondly, whether the bases entering into the judgment were common for the entire group of children studied. Unfortunately, agen-

cy records do not reveal in any consistent manner the bases for their decisions. Hence, the inclusion of children placed at older ages would introduce serious disturbing factors in a research population. Further, any research on adopted children placed at older ages involves the measurement of the influence of the environment previous to the adoptive or foster one under consideration. Because of these reasons it was clear that our experimental population must be composed of children placed in their adoptive homes at as near zero age as possible. Only with such a population could we hope to secure random placement of children and thus reduce the operation of the unmeasurable influence of selective placement. It was recognized, nevertheless, that judgments of intellectual promise on the basis of family history still remained. The ideal experimental population would include only those children whose age at the time of placement is so young as to preclude prediction of future mental development and for whom no evidence of family history is available to agencies or persons placing the children. The complete elimination of family history is impossible, since the number of foundlings is relatively small. However, when the complexity of human inheritance is considered, as well as the infinite variety of factors that enter into the determination of socio-economic status in a competitive society, it is clear that inferences as to the mental ability of an individual from isolated facts of family history are highly unreliable. General trends, however, in mass evidence have been found to be consistent. For example, the progeny of college-trained pro-

fessional people are generally above the average in mental ability, while the children of the unschooled laborer are somewhat inferior. A judgment as to the intelligence of a particular infant from a knowledge of his family background nevertheless would be very unreliable. The child may fall far below or far above the average of his parental group. Yet if child placing agencies pursued a consistent policy of relating cultural status of background to that of the adoptive home in all placements, a definite resemblance between foster parents and children would result. The contribution of family history to a judgment of the intellectual level of our experimental children and its consequent weighting of the observed resemblance between adoptive parent and child will be discussed in Section VII.

The imperative need of a check upon our methods committed us to a control group of true parents and offspring from the very inception of the study. Only from a study of a group of children who had been given the same tests and measures as the adopted children could we hope to attach any meaning to the results observed in the adoptive group. High or low correlations between adoptive parents and children, for example, might well be said to be the result of factors peculiar to our measuring instruments. The behavior of any particular group of human beings has meaning only in so far as we know how human beings in general behave. With measurable environment identical for both groups of children, differences in the relationship of child's intelligence to environment obviously must be the result of the presence of a common hered-

ity in the case of true parents and offspring and the result of the absence of hereditary likeness in the case of adopted parents and their children.

III

SUBJECTS

Source of Subjects

Once the decision was made to limit the investigation to children placed in their adoptive homes at a very early age, it was apparent that the records of adopted children deposited in the Children's Bureau of the State Board of Control would be our most complete source for subjects, since the adoption records of the entire state are available there.

Due to the relatively small number of legitimate children that are available for adoption in early infancy, it was deemed better to limit our subjects to illegitimate children. An additional point of significance in favor of illegitimate children was the greater probability of securing a population whose intelligence would be normally distributed. Legitimate children are ordinarily available for adoption only because of serious intellectual and economic inadequacies in their parents or immediate relatives. Illegitimate children, on the other hand, are relinquished for many reasons, namely: the youth of the parents, the social stigma attached to the illegitimacy situation, economic inadequacy of parents, and intellectual incompetency of the parents. The economic inadequacy of unmarried parents is frequently associated with youth, while the economic inadequacy of married parents generally arises from intellectual and personality deficiencies. In general, a population of legitimate dependents appears to come from a narrow socio-economic range, while il-

legitimate dependents come from a more variable family background.

Our first step, then, in anticipation of our research project was the tabulation of the factual items of family history for the illegitimate children adopted in Minnesota during the period 1918-1928. This period was chosen because it would provide children who would be not less than 5 nor more than 14 years of age at the time of the field investigation, 1932-1933. Records were available for 2449 children. Our transcript included information on the personal history of the child, the true parents and the foster parents.

Experimental Group, Criteria of Selection

In order that the least possible ambiguity exist in our results, the experimental group was limited to:

1. *Children placed in their adoptive homes at the age of 6 months or younger.* (The mean age of placement was 2.5 months) At this early age precise judgments of mental ability on the basis of test performance, physical development, or overt behavior are highly improbable. Further, this criterion assures from early infancy an environment that is no more or less changing in character than that enjoyed by children in general. Moreover, it definitely avoids the difficulties which would arise in attempting to measure the influence of environment previous to the adoptive one under consideration.

2. *Only those adopted children who were known to be of white race, non-Jewish, north-European ex-*

traction. This prerequisite tends to reduce the possibilities of a fortuitous resemblance between adoptive parent and child on the basis of racial regression. In addition it minimizes the possibility of a spurious heterogeneity arising from uncontrolled factors relating to race. Further, it limits the group to one which is similar in composition to the one on which the Stanford-Binet test was standardized.

3. *Children who were not less than 5 nor more than 14 years of age at the time of investigation.* This age range is conceded to give the most reliable test results.

4. *Children reared in communities of 1000 or more.* In this way we attempted to equalize the influence of such environmental factors as churches, clubs, and schools. No farm children are included. Ninety-five per cent of the group have been reared in communities of over 10,000.

5. *Children who were legally adopted by married persons.* Thus we secured a group where the legal relationship and responsibility between parent and child was the same as that of true parent and offspring.

6. *Adoptive parents who were of white race, non-Jewish, north-European extraction.* With this criterion we attempted to reduce the possibility of adventitious resemblance and further reduced the possibility of securing non-English-speaking homes.

By adhering rigidly to the foregoing criteria it is believed that we have controlled the element of selective placement to a point beyond the facilities of earlier

investigators and to the highest possible degree that present day child adoption permits. Fitting the child to the home on the basis of coloring, physique, and religious faith, all of which occur, could hardly give rise to mental resemblance. Selective placement upon the basis of cultural status, however, is still possible. But since the preadoptive records did not reveal the facts on this point, we can only infer its existence or nonexistence from an analysis of the relationship of certain indices of cultural status. Evidence relative to this point will be presented in Section VII.

In our earliest considerations of a population we conceived a research group which would sample the population of adoptive homes distributed from a socio-economic standpoint as male occupations are distributed in the general population. Because of the limited number of children placed in homes of the laboring class this plan had to be abandoned. In its place we accepted all children available in the two lowest occupational groups and secured at least 40 children at every other level. With these numbers we have not only obtained a fair picture of environmental differences contingent on occupational status, but have also secured a fair sample of the selective placement that may operate on the basis of cultural status. A small number at any level might give a distinctly distorted picture.

Control Group, Criteria of Selection

With the primary purpose of a control group serving as a check upon the validity of our methods, each

adopted child was matched with an *own* child as follows:

1. *For sex.*

2. *Within an age range of plus or minus 6 months.*

3. *Whose fathers' occupations fell in the same group on the Minnesota Occupational Scale.*

4. *Whose fathers' school attainments agreed within plus or minus one school grade level.*

5. *Whose mothers' school attainments agreed within plus or minus one school grade level.*

6. *Whose parents were white race, non-Jewish, north-European extraction.*

7. *Whose residence has been in communities of 1000 or more.*

The problem of matching cases was very arduous and time consuming. The degree to which we have been successful in locating two groups of children living under similar environmental conditions: one, unrelated by blood to the adults rearing them, the other, the offspring of the adults shaping the environment, will be shown by a series of tables and graphs in Section V.

In matching cases for occupation and education, we employed the two most objective indices of cultural status that are available. In 12 cases the educational criterion was not adhered to. It was necessary to be content with agreement in education for one parent. In these cases, however, the educational disparity between the other adoptive parent and his control was

held within the ordinary school groups of elementary, high school, or college level.

A typical match is illustrated by the following example. A lawyer, in the person of an adoptive father who had completed college and whose wife had finished the eleventh grade in high school, might be matched with an electrical engineer of not less than three or more than five years of college and whose wife had completed at least the tenth, but not more than the twelfth grade in high school, provided the sex and age of their respective children agreed.

If environment is dominant, it would seem that the trend of any trait concerning the children and their environment would be similar in direction and magnitude for both the adopted and control populations. Certainly, our adopted children should clearly reveal the relationship of environment to attributes which are not reciprocally affected by the innate tendencies of the child. For example, the occupation of the adoptive father is obviously not a function of the intelligence of the adopted child, while the number of children's books in the adoptive home and the intelligence of the child are reciprocally dependent. The books may be in the home because intelligent children enjoy books. Or the children may be more responsive and alert because the books are in the home. When the age of our children at the time of the test is considered, it is apparent that there are many factors in the adoptive home whose existence is entirely independent of the child and, therefore, whatever relationship exists between these factors and the child may be regarded as

a measure of the influence of environment. The relationship between parent and true offspring, however, is a complex of environment and heredity. Here, for example, the child's intelligence may be not only the result of the quality of the environment that the parents have provided, but also an inherited characteristic from the parents. From the adopted population we should be able to get a measure of the influence of environment; from the control population a measure of the influence of environment reenforced by heredity. Whatever unreliability exists because of imperfections in our measuring instruments will be similarly existent in both populations. Further, whatever the accumulated effect of environment may be, it will be operative in both populations in the same direction since both have enjoyed what might be termed an ordinarily continuous environment.

IV

METHOD

Approach to Subjects

With but few exceptions each experimental and control family was interviewed three times. The first interview was concerned with an explanation of the purpose of the study and with the collection of social data. At that time the mother was ordinarily seen alone when the child was away at school.

The second interview was devoted to testing the child. This was done either in the home or at the University. The former was allowed for all cases located outside of Minneapolis and St. Paul and for residents of the latter communities when conditions made a trip to the University impossible. Every attempt was made in the case of the latter to approximate good laboratory standards for individual mental examinations. With but few exceptions the entire control group of children were examined at the schools which they attended. In their situation the study could be explained to the school authorities since there was no danger of disclosing confidential information.

The third interview was given over to a report of the child's test performance followed by the administration of a set of intelligence tests to the parents. This interview was usually held in the evening when both parents could be present. In a small number of cases, the second and third interviews were combined; while one investigator tested the child, the other tested the parents.

The time required to carry out our program with an individual family varied from five to seven hours. Only rarely was it completed in less than five hours.

The confidential character of our knowledge of the presence of an adopted child in a home demanded that we proceed with caution. Hence the introduction of the study to a parent followed a pattern designed to respect this confidence, as well as to secure the highest possible cooperation. First, a letter which explained the character and extent of the investigation was dispatched. It contained no reference to adopted children.

> The Institute of Child Welfare of the University of Minnesota is making a research study of the child in his home environment. Four hundred homes located in various parts of the state of Minnesota have been carefully selected to take part in this study. Your home was among those chosen.
>
> Other than a total of about two or three hours of time no demands are made of parents. A field worker will want to see you and then give your child certain simple ability and interest tests. Most children regard the tests as games and seem to enjoy them thoroughly. The tests will be given in your home or at the University, whichever seems the more convenient. The study is entirely non-commercial. The fullest practical and scientific value of the study will be possible only through the assistance of parents. Your cooperation will be a contribution to an important scientific project.
>
> During the progress of the study parents will be given the results of their children's tests. These should be of considerable interest and value. They will be used for statistical purposes only and without the names of the individuals participating in the study.

Gauging the receipt of the letter and allowing sufficient time for both parents to have perused its contents, inquiry was made by telephone as to whether or not the communication had been received. If the reply was in the affirmative, an opportunity to visit the home in order to explain the study was promptly requested. If the letter had not been received, the study was briefly described and a future communication promised. Reference to the status of the child was avoided in all telephone communications. However, a number of persons (30-35) asked precipitously whether the presence of an adopted child in their home had been the basis of selection. Admitting the inclusion of *adopted* as well as *own* children in the study, a direct answer to the question was usually evaded—not of course with invariable success. However, in most cases we were able to postpone the discussion of the exact basis of selection until the home visit. At that time the basis of selection of subjects, the importance of adopted children in a study such as we were undertaking, and a complete description of our plan was frankly discussed. Even in families where the child had not been told of his adoption and the parents had at first feared that cooperating in our research would disclose the fact, the explanation of the significance of the study along with the assurance that the identity of the individual would be held in strict confidence was sufficient to secure excellent response. For a small number of families who could not be reached directly by telephone, the original announcement letter indicated possible visiting hours. A self-addressed card

on which to record their preference was also enclosed.

One hundred and ninety-four, 84 per cent of the total number (231) of adoptive families asked to cooperate, participated in the study. Thirty-seven families, 16 per cent, refused their cooperation. Twenty-three of these were telephone refusals. Only 14 refused after a personal interview. Although the telephone refusals constitute 62 per cent of the total group refusing, the method was distinctly economical of time. Further, suspicion and antagonism would in all probability have been greater had an investigator appeared unannounced at the home. As evidence of the authenticity of the investigation, the parents had before them a letter on official stationery, signed by the Director of the Institute of Child Welfare.

The reasons offered by those parents who refused their cooperation were the following: a fear that the knowledge of adoption would reach the child, child sensitive about adoption, uninterested in study, and unwilling to have child tested. In the group that cooperated, only 50 per cent of the children had been told that they were adopted. The proportion of children without instruction as to their true status among the refusals is probably greater. Unwillingness to have child tested was given in only three cases. A larger number in this category would probably seriously affect our results. A complete analysis of the entire group that refused cooperation and a discussion of the probable effect of their refusal on our results will be presented in Section VII.

The approach to our control families followed that

used for the adoptive homes. The same announcement letter was used. Due to errors in matching cases, 40 control families over and above those used in our analysis were given partial or complete study. Here the number of refusals was small. Because they could always be replaced by other families that resembled the adoptive family equally as well and who would presumably equally well expose familial resemblance, no description was compiled.

Tests Employed and Their Administration

The tests and measures used for both the adoptive and control families were the following:

1. A set of three blanks entitled *The Child and His Environment* covered the family and personal history of the child. Space was provided for specific information as to the identity of the child, the condition of his health, the cultural background of his true parents and for the replies to 88 questions relative to the cultural, economic and social status of the adoptive home. This last category of information constitutes our quantitative measure for environment. All replies were directly recorded in the mother's presence. The questions pertaining to the true parents were almost invariably held over to the latter part of the interview. Occasionally, however, the adoptive parent volunteered the information in advance of any specific questions. None gave the impression of withholding information on these points, although the majority were unable to answer the questions pertaining to the child's family

background. Obviously this section was not pertinent in the case of the control children.

The second blank permitted a description of the home as gleaned from informal conversation with the mother. The child's personality, behavior, his school progress, and a narrative of the circumstances surrounding his selection were the topics usually discussed. These data were recorded subsequent to the interview, but as soon after as possible.

The third schedule designed to contribute to our knowledge of family and personal history was the interest interview with the child. This was usually administered after the mental examination and covered the child's educational and social activities. The questions were asked directly and the replies were recorded immediately.

2. The Stanford Revision of the Binet-Simon Tests as described in Terman's "The Measurement of Intelligence" was administered to all the children. The procedure recommended by its author was rigidly adhered to, i.e., the testing was carried down to a level at which all tests were passed and up to a level at which all tests were failed. For those children who exceeded the limits of the test the correction worked out in connection with the Stanford Study of gifted children was applied. (Terman, 8, p. 42) In order to minimize errors contingent on the personal differences of examiners, a single examiner administered all of the tests to the adopted children.[1] Similarly in the case of the con-

[1]Miss Winona Morgan, the chief assistant, tested every adopted child. Miss Amanda Herring tested all of the control children. Both examiners had M.A. degrees in psychology and both were trained in mental testing by Dr. F. L. Goodenough.

trol children a single examiner was used. All tests were checked and rescored by the mental test division of the Institute of Child Welfare.

3. The Woodworth-Mathews "Personal Data Sheet," a questionnaire of 75 questions designed to reveal psychotic tendencies, was given to all children age 10 years and over. Two questions (44 and 45) which pertain to adopted children were blotted from the booklets. According to the author the total number of unfavorable responses provides an index of an individual's emotional stability. Since all other variables entering into our analysis are expressed by increments which increase with favorableness, the Woodworth-Mathews questionnaire was scored for total favorable responses. No claims are made for the validity or reliability of the questionnaire. It is simple to administer and supplies a ready instrument for comparative analysis.

4. The Otis Self-Administering Test of Mental Ability—Intermediate Form A, was given to the parents. The Intermediate examination, designed for grades 4 to 9, was chosen instead of the higher examination designed for high school and college students because of the great difficulty exhibited by 20 parents who tried the higher examination. They found the illustrative questions imposing and in general their "set" to the test situation was one of anxiety. And although our social data subsequently showed the mean schooling of the adoptive parents to be eleventh grade and that of the control parents, tenth grade, the average older adult is unfamiliar with the test situation and

"shys off" anything of this nature that appears difficult. The fact that the test results were to be used primarily for comparative purposes further justified the use of the simpler form.

In order to reduce the fear that might arise from the word *"test"* the title at the top of each Otis examination booklet was pasted over with a piece of paper resembling in texture and weight that of the booklet. In referring to the test the investigator called it a list of questions. Otherwise its administration adhered to the directions given in the manual. Twenty minutes was allowed for its execution. As in the case of the children's mental tests, each Otis examination was checked for scoring errors. In choosing the Otis the author was influenced not only by the ease with which it can be administered, but also because it would provide a check on Freeman's results since he had used it in his study of the influence of environment on intelligence.

5. The Stanford-Binet Vocabulary Test was given to the parents. Total vocabulary score was calculated on the basis of the subject's responses to the first list of words. Definitions were carefully checked according to the directions in Terman, *The Measurement of Intelligence.* This test was included in our program because of its high correlation with the whole Binet scale and, secondly, because it had been used by Burks in her study of nature-nurture and intelligence.

6. A transcript of the social agency's record of the child and his true parents was made. The items of education and occupation of forbears were rechecked with the original record in order to have available the most reliable information possible on the cultural level of the child's background.

V

DESCRIPTION OF EXPERIMENTAL GROUPS

Our Adopted and Control groups will be described by a series of tables and graphs. The first of the series concerns the children—their age, school grade, and intelligence.

Age

In Table 1 and Figure 1 age is shown. From these figures, it is apparent that this research concerns two groups of children who are practically identical in age; the mean for the Adopted group is 9.3±2.5 years, for the Controls, 9.4±2.5 years.

TABLE 1

Age Distribution of Adopted and Control Children

Age in years	Adopted children	Control children
15	1	4
14	10	7
13	14	17
12	18	16
11	23	23
10	24	24
9	20	22
8	23	28
7	35	33
6	20	15
5	6	5
M	9.3	9.4
SD	2.5	2.5
N	194	194

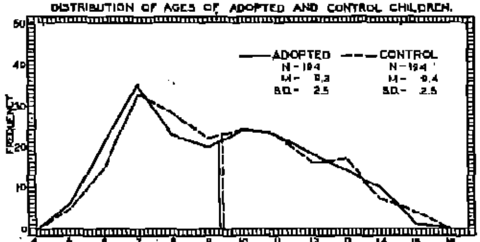


FIGURE 1

SCHOOL GRADE

Although the children were not matched for school grade, the agreement in school attainment as shown in Table 2 and Figure 2 is striking. The mean school

TABLE 2

DISTRIBUTION OF SCHOOL GRADE OF ADOPTED AND CONTROL CHILDREN

School grade		Adopted children	Control children
10		—	4
9		10	5
8		9	8
7		16	17
6		17	19
5		22	23
4		26	23
3		22	27
2		28	25
1		33	31
Kgn.		8	11
	M	3.9	3.9
	SD	2.5	2.5
	N	191	193

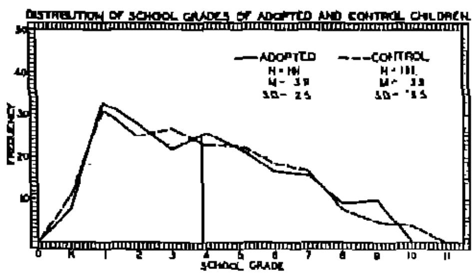


FIGURE 2

grade of the Adopteds is 3.9 with a standard deviation of 2.5; the mean for the Controls is also 3.9 with a standard deviation of 2.5. Hence, whatever may be the influence of environment functioning through the institution of the school, we should expect it to affect the two groups equally.

Intelligence

The distribution of intelligence test scores is shown in Table 3 and Figure 3. The average IQ of the

TABLE 3

Distribution of IQ of Adopted and Control Children

Stanford-Binet IQ	Adopted children	Control chldren
160-164	0	2
155-159	1	1
150-154	0	0
145-149	1	2
140-144	2	0
135-139	1	6
130-134	6	7
125-129	11	19
120-124	26	13
115-119	21	19
110-114	31	23
105-109	33	21
100-104	23	26
95- 99	21	24
90- 94	6	17
85- 89	10	5
80- 84	0	7
75- 79	1	2
M	110.5	109.7
SD	12.5	15.4
N	194	194

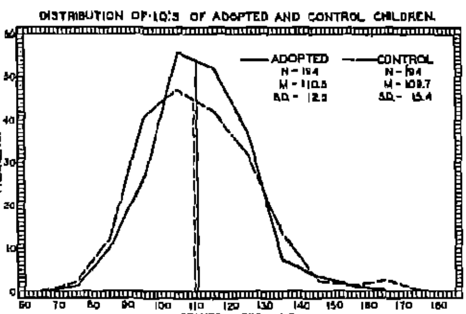


FIGURE 3

Adopted children, as determined by the Stanford-Binet Test, is 110.5, of the Controls 109.7. A statistically

insignificant difference of .8 exists between the two groups.

Although practical identity exists in test intelligence, a consideration of why the averages of the two groups are above those of children in the general population is pertinent. Since it has been demonstrated that girls do better on verbal tests than boys, can the disproportionately larger number of girls in these data, 61% as against 39% boys, have determined these averages? A separate analysis of the test performance of the boys and girls in both populations showed the difference between the sexes to be insignificant. The mean IQ of adopted boys is 109.6, of adopted girls, 111.4. The ratio of the difference to its standard error is 1.02. In the case of the control children, the mean for boys is 110.0, and for the girls, 109.5. Here the ratio of the difference to its standard error is .004.

The next consideration was selection. The probability that it is operative in the Control population is reasonable, since in matching cases on the basis of adoptive father's occupation the number of homes in the upper economic levels outweighs the number of homes in the lower occupational groups. In fact, there are five times as many homes of the professional class as this level appears in the general population, and the excess of homes from the business managerial group is two times their occurrence in the general population, while the proportion of homes in the two lowest occupational groups is about one-half their frequency in the general population. Hence the positive deviation of average IQ is to be expected, since here we have se-

lective heredity and superior environment, both operative.

A further condition which would tend to lift the intellectual level of control children in the lower occupational brackets, and hence raise the general average of the entire group, is the fact that control parents were selected on the basis of their agreement with adoptive parents in school attainment. When this criterion is introduced into groups whose occupation is not dependent upon amount of schooling, as in the case of the unskilled workman, the intellectual level of the group is thereby raised.

What is the evidence of selection in the Adopted group? If we have random placement of adopted children, an excess of homes from the upper economic levels should not have the effect observed, unless it can be ascribed to the influence of environment. Before such a conclusion is warranted, the elimination of unpromising candidates for adoption must be considered. Inquiry into the laws concerning the adoption of children in Minnesota reveals that, since 1917, annulment of adoption is possible if the child develops to be feebleminded, epileptic, insane, or afflicted with venereal disease, as a result of conditions existing prior to the adoption. Annulment because of feeblemindedness occurred in one case that we attempted to locate. How many children were placed out but not adopted because of mental defect we have no way of knowing. Since 1917 a trial period of six months' residence in the home before adoption has been required by law. All of our children have resided five or more years

and all had been legally adopted. Further, since 1925, a child has been regarded as unsuitable for adoption, if one or more of the following obtain: (1) the offspring of incestuous cohabitation, (2) one or both parents feebleminded or insane and the mental state of the child is as yet undetermined, (3) if the child is crippled or deformed or afflicted with tuberculosis, venereal disease, or other contagious or offensive disease that render his presence a menace to others.

The restriction denying adoption to children of feebleminded persons until the offspring's intelligence is determined would alone be sufficient to raise the general intellectual level of adopted children. Since all of our children were placed and adopted previous to an age at which reliable tests of mental ability can be made, it follows that there was no known history of feeblemindedness in their immediate background.

On the basis of the foregoing conditions, we should expect positive skewness in the intelligence level of our adopted children. It should also be noted that in an analysis of the cultural background of 10,000 potential dependent children made by the author and published in *The Journal of Genetic Psychology,* December, 1932, it was found that dependent illegitimate children who were relinquished for adoption had more superior family histories judging from occupation and education of parents than those children retained by their true mothers. Hence, the element of selected heredity appears to be playing a distinct part in the positive deviation of the average IQ of our Adopted children. The part that environment may be playing will be discussed in Section VI.

The environmental indices of our populations include: educational attainment, test intelligence scores, and vocabulary scores for both parents, the occupational status of the fathers and home rating as expressed on a quantitative scale designed to measure urban home conditions.

Education of Parents

Tables 4 and 5 and Figures 4 and 5 describe the school grade attainment of the two sets of parents. The agreement is marked. The mean in the case of the fathers of Adopted children is 11.2±4.3, of Controls, 11.4±4.0. The mean school grade attainment of the

TABLE 4

Distribution of Education of Adoptive and Control Fathers

School grade completed		Adoptive fathers	Control fathers
21		0	1
20		1	3
19		4	5
18		13	2
17		5	7
16		34	30
15		2	9
14		6	9
13		1	6
12		24	21
11		6	5
10		12	14
9		7	11
8		46	46
7		11	9
6		10	8
5		2	5
4		2	1
3		6	1
2		0	0
1		1	0
	M	11.2	11.4
	SD	4.3	4.0
	N	193	193

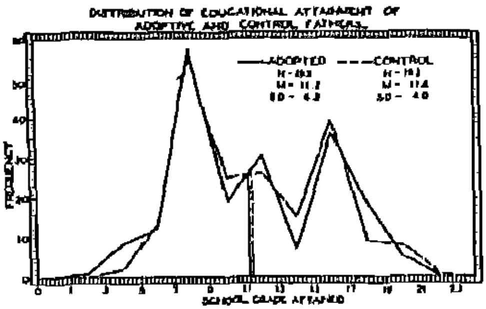


FIGURE 4

TABLE 5

DISTRIBUTION OF EDUCATION OF ADOPTIVE AND CONTROL MOTHERS

School grade completed		Adoptive mothers	Control mothers
17		0	1
16		20	20
15		4	10
14		11	7
13		6	11
12		45	43
11		11	12
10		15	18
9		8	14
8		50	40
7		6	6
6		5	3
5		7	7
4		3	2
3		1	0
	M	10.5	10.8
	SD	3.1	3.0
	N	192	194

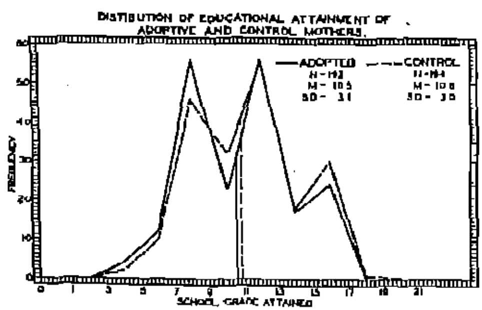


FIGURE 5

mothers of Adopted children is 10.5±3.1, of Control mothers 10.8±3.0.

INTELLIGENCE OF PARENTS

The similarity in intelligence test performance of the Adoptive and Control parents is notable. As shown in Table 6 and Figure 6, the mid-parent score for

TABLE 6
DISTRIBUTION OF MID-PARENT OTIS SCORE OF ADOPTIVE AND CONTROL PARENTS

Otis score		Adoptive parents	Control parents
70-74		4	10
65-69		24	24
60-64		24	28
55-59		28	26
50-54		16	12
45-49		27	18
40-44		16	12
35-39		12	12
30-34		9	10
25-29		8	4
20-24		5	1
15-19		1	2
10-14		3	1
0- 9		0	0
	M	50.0	52.0
	SD	13.6	13.1
	N	177	173

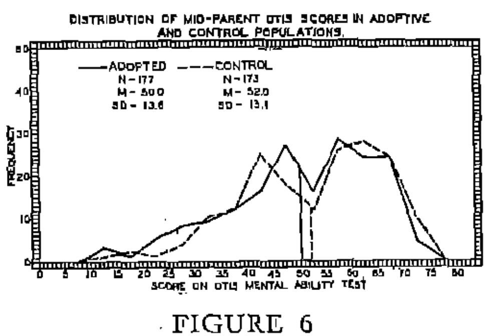


FIGURE 6

Adoptive parents on the Otis test of mental ability is 50.0±13.6, for the Control parent it is 52.0±13.1.

VOCABULARY ABILITY OF PARENTS

On the Stanford-Binet vocabulary test the difference in the performance of the two groups of parents is insignificant. The average for the Adoptive parent is 65.0±12.0, the average for the Control parents is 63.2±12.1 as given in Table 7 and Figure 7.

TABLE 7
DISTRIBUTION OF MID-PARENT VOCABULARY SCORE OF ADOPTIVE AND CONTROL PARENTS

S.B. Vocabulary score		Adoptive parents	Control parents
85-89		3	6
80-84		19	8
75-79		16	17
70-74		36	28
65-69		31	21
60-64		26	16
55-59		15	27
50-54		15	21
45-49		4	9
40-44		2	7
35-39		2	4
30-34		4	0
25-29		0	0
20-24		1	0
	M	65.9	63.2
	SD	12.0	12.1
	N	174	164

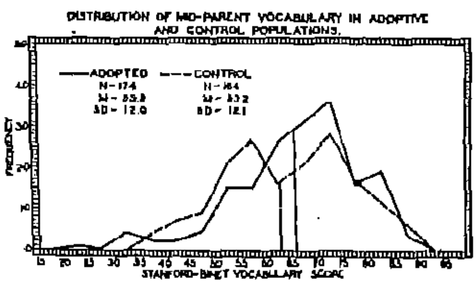


FIGURE 7

Occupational Status

Table 8 and Figure 8 show the agreement in home environment as determined by the occupational status of the fathers of our Adopted and Control children.

TABLE 8

Occupational Classification of Adoptive and Control Fathers

Occupational group	Adoptive fathers	Control fathers
I Professional	41	41
II Business manager	40	40
III Skilled trades	44	44
IV Farmers	0	0
V Semi-skilled	45	45
VI Slightly skilled	20	20
VII Day labor	4	4
M	3.1	3.1
SD	1.3	1.3
N	194	194

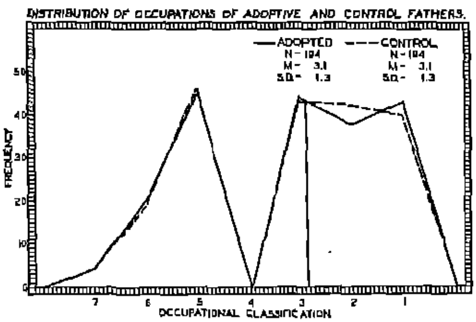


FIGURE 8

Environmental Status

In environmental status score, Table 9 and Figure 9, the adoptive homes show a significant though not very large excess over the control homes, *137.9±54.3* as against 118.7±59.6. Since this measure is a combined expression of occupational status, education of parents,

TABLE 9
DISTRIBUTION OF ENVIRONMENTAL STATUS SCORE OF ADOPTIVE AND CONTROL HOMES

Environmental status score		Adoptive homes	Control homes
225-239		8	0
210-224		18	8
195-209		15	13
180-194		16	21
165-179		12	10
150-164		15	19
135-149		13	16
120-134		20	13
105-119		20	9
90-104		13	13
75- 89		12	16
60- 74		22	20
45- 59		7	11
30- 44		2	11
15- 29		1	8
0- 14		0	6
	M	137.9	118.7
	SD	54.3	59.6
	N	194	194

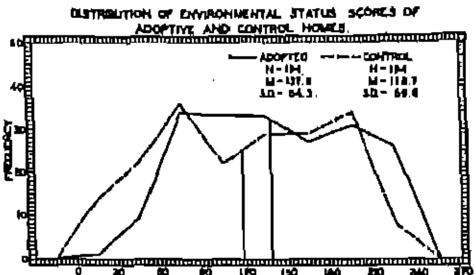


FIGURE 9

economic status, degree of social participation, cultural materials, and child training facilities in the home, and since our Adoptive and Control parents are almost identical in occupation and education, the difference must be in the other factors. The smaller size of the adoptive family no doubt permits greater material

possessions as well as freedom for a greater amount of participation in community activities.

In general it appears that the environment in the Control group is truly comparable to that of the Adopted group. To the extent that the foregoing indices signify similarity in environmental potency the homes of the two groups of children may be regarded as equally stimulating. Conceding our inability to obtain a measure of the stimulating or depressive effect of any one environment for any one child, the general influence in the two groups would be most probably the same. Hence, if environment is dominant, we should expect the magnitude of the relationship between child intelligence and any single index or any combination of indices to be relatively the same for both groups of children.

VI

MAIN RESULTS OF THE STUDY

The relationship between test intelligence of children and various attributes of their home environment is shown in Table 10. Since intelligence and age of

TABLE 10

CHILD'S IQ CORRELATED WITH OTHER FACTORS

Correlated factor	Adopted children *r*	P.E.	N	Control children *r*	P.E.	N
Father's Otis score	.15	.05	178	.51	.04	175
Mother's Otis score	.20	.05	186	.51	.04	191
Mid-parent Otis score	.18	.05	177	.60	.03	173
Father's S.B. vocabulary	.22	.05	177	.47	.04	168
Mother's S.B. vocabulary	.20	.05	185	.49	.04	190
Mid-parent S.B. vocabulary	.24	.05	174	.56	.03	164
Environmental status score	.19	.05	194	.53	.03	194
Cultural index of home	.21	.05	194	.51	.04	194
Child training index of home	.18	.05	194	.52	.04	194
Economic index of home	.12	.05	194	.37	.04	194
Sociality index of home	.11	.05	194	.42	.04	194
Father's education	.16	.05	193	.48	.04	193
Mother's education	.21	.05	192	.50	.04	194
Mid-parent education	.20	.05	193	.54	.03	194
Father's occupational status	.12	.05	194	.45	.04	194

child have been demonstrated to be negatively correlated (in these data, age and IQ for adopteds correlated from —.17 to —.19, for controls from —.13 to —.18), age has been partialled out and the relationships are expressed in product moment correlations. Because it was not possible to obtain full information for all the persons participating in the study, the number of cases varies for each correlation.

Although the difference between corresponding correlation coefficients in the Adopted and Control group is consistent and striking, their comparability must be determined before any interpretations are made. The

test of comparability is equal variability. A reexamination of the data in Section V shows almost perfect agreement in the variability of environmental factors entering our correlational table. Equal variability does not exist for test intelligence, however, in the two sets of data. In the case of the Adopted children it is 12.5, for the Control children, it is 15.4. Although the difference is not large, correction should be made if two equally comparable series of coefficients are desired. Since the nature of the curtailment is known and exists in only one trait, the correction evolved by Pearson may be applied.[2] The corrected correlations are presented in Table 11.

TABLE 11

CHILD'S IQ CORRELATED WITH OTHER FACTORS

(r corrected for unequal range in child's IQ)

Correlated factor	Adopted children			Control children		
	r	P.E.	N	r	P.E.	N
Father's Otis score	.19	.06	178	.51	.04	175
Mother's Otis score	.24	.06	186	.51	.04	191
Mid-parent Otis score	.21	.06	177	.60	.03	173
Father's S.B. vocabulary	.26	.06	177	.47	.04	168
Mother's S.B. vocabulary	.24	.06	185	.49	.04	190
Mid-parent S.B. vocabulary	.29	.06	174	.56	.03	164
Environmental status score	.23	.06	194	.53	.03	194
Cultural index of home	.26	.06	194	.51	.04	194
Child training index	.22	.06	194	.52	.04	194
Economic index	.15	.06	194	.37	.04	194
Sociality index	.13	.06	194	.42	.04	194
Father's education	.19	.06	193	.48	.04	193
Mother's education	.25	.06	192	.50	.04	194
Mid-parent education	.24	.06	193	.54	.03	194
Father's occupational status	.14	.06	194	.45	.04	194

Despite the severity of the correction the absolute change in magnitude of our correlations is not great. The greatest single increase is .05; on the average the

[2]From Kelly (4, pp. 225, 316).

correlations are increased .038 points. Note that the difference between corresponding coefficients in the Adopted and Control series continues. For the Adopted children they are consistently low, about .20. In the Control group they maintain the level usually found for hereditary physical characteristics, .50. In the case of the latter group heredity and environment are both operative. Hence variance in intelligence is accounted for by variance in heredity and environment combined to the extent of about 25 per cent (square of r .50). In the Adopted group, however, where environment is functioning independently of heredity, variance in intelligence is accounted for by variance in environment only to the extent of about 4 per cent (square of r. 20). If we neglect whatever artificial heredity selective placement of adopted children may have introduced into the data, these coefficients are clear evidence of maximum variance in intelligence with variance in environment. Apparently environment cannot compensate for the lack of blood relationship in creating mental resemblance between parent and child. Heredity persists.

A second type of analysis of our data appears in Table 12, where the mean intelligence quotient of Adopted children in each successive occupation level is compared with the intelligence quotient of Control children similarily classified according to occupation of father. Note the constancy of the IQ of Adopted children, irrespective of occupational level. Its progression is insignificant. When variability in IQ within each occupational group is considered the children

TABLE 12

COMPARATIVE ANALYSIS OF INTELLIGENCE OF ADOPTED AND CONTROL CHILDREN AND ENVIRONMENTAL STATUS OF HOMES CLASSIFIED ACCORDING TO OCCUPATION OF FATHER

	Adopted children					Control children				
		Intelligence quotient		Environmental status			Intelligence quotient		Environmental status	
Occupation of father	N	M	S.D.	M	S.D.	N	M	S.D.	M	S.D.
I Professional	43	112.6	11.8	194.6	27.2	40	118.6	12.6	180.4	29.1
II Business manager	38	111.6	10.9	171.3	40.2	42	117.6	15.6	160.7	31.1
III Skilled trades	44	110.6	14.2	133.2	35.2	43	106.9	14.3	106.3	43.4
IV Farmers	—	—	—	—	—	—	—	—	—	—
V Semi-skilled	45	109.4	11.8	94.0	30.3	46	101.1	12.5	77.6	37.4
VI Slightly skilled and VII Day labor	24	107.8	13.6	74.7	28.7	23	102.1	11.0	40.1	26.7

in the lowest level almost completely overlap the children in the highest group. The same is true when occupational groups V and I are compared. The difference is entirely effaced between occupational classes III and I. If we ignore the very lowest occupational bracket (VI and VII) in which the number of cases is considerably less than in the other levels, a difference of only one IQ increment exists between the successive occupational classes of Adopted children.

The Control children, on the other hand, advance conspicuously in mean level of intelligence with fathers' occupation. The difference in IQ between the lowest occupational levels and the middle group (III, skilled workmen, clerks, etc.) is as great as the difference in IQ between the lowest and highest occupational group of the Adopted children. Although the children in the two highest occupational classes (business managerial and professional) are undifferentiated they are widely separated from the children of the middle group (about 12 IQ points). The absolute difference in child's IQ between the extreme occupational levels in the Control group is three times as great as the difference between the extreme levels of Adopted children. The fact that the children of each occupational group are almost identical in age should be borne in mind. If the children in the highest occupational levels were younger than those in the middle and lowest groups, then their superior rating in IQ might be said to be a function of age. It will also be recalled that each Adopted child was matched with a Control child of the same age and whose father's occupation

was in agreement with that of the adoptive fathers. Hence, cross-comparisons are entirely valid.

The probability of differences in IQ continuing in the same direction with occupational status in the case of similarily chosen populations as those observed here is shown in Table 13. Apparently none of the differ-

TABLE 13

COMPARATIVE ANALYSIS OF THE PROBABILITY OF DIFFERENCES IN IQ OF CHILDREN EXPRESSED BY $D/\sigma_{diff.}$ CONTINUING IN THE SAME DIRECTION AS IN THE EXPERIMENTAL POPULATIONS CLASSIFIED ACCORDING TO FATHER'S OCCUPATION

Occupational groups	Adopted children II	III	V	VI & VII	Control children II	III	V	VI & VII
I	.40	.72	1.27	1.45	.32	3.96	6.45	5.43
II		.36	.88	1.15		3.29	5.44	4.66
III			.43	.80			2.03	1.52
V				.49				.34

ences for the Adopteds are reliable. Reversal of direction might occur in another population; while in the Control population the differences between the two highest groups and every other group are clear and dependable. The probability of a difference in the same direction between the middle and the lowest level is greater for the Controls than the probability of the recurrence of any single difference in the Adopted population. Clearly, environment as typified by occupational status does not compensate for the absence of blood relationship between parent and child.

The foregoing observations take on added significance when the environmental status scores of the successive occupational levels are studied. We note that

in both populations environmental status score increases with occupation and further that the magnitude of each successive difference is more than 2.6 times its standard error (Table 14). It should also be noted

TABLE 14
COMPARATIVE ANALYSIS OF THE PROBABILITY OF DIFFERENCES IN ENVIRONMENTAL STATUS SCORES EXPRESSED BY $D/\sigma_{diff.}$, CONTINUING IN THE SAME DIRECTION AS IN THE EXPERIMENTAL HOMES CLASSIFIED ACCORDING TO FATHER'S OCCUPATION

Occupational groups	Adoptive homes II	III	V	VI & VII	Control homes II	III	V	VI & VII
I	3.02	9.12	16.40	16.70	2.96	9.19	14.32	19.43
II		4.53	9.74	11.02		6.65	11.37	16.41
III			5.63	7.40			3.33	7.65
V				2.61				4.79

that at all levels the mean score of the Adoptive homes is higher than that of Control homes. If intelligence progresses with environment independently of heredity then as great increases in IQ should be expected in the Adopted group as in the Control group.

Considering test intelligence and vocabulary scores of parents, measures which may be more reflective of innate capacity than environmental status score, the same tendency of progression with occupation is noted for both Adoptive and Control parents. (Table 15). In the case of Control parents the occupational levels appear about equally spaced in respect to test intelligence and vocabulary scores. In the Adoptives the two lowest levels are undifferentiated in score as are also the two highest. Yet both extremes are equally spaced from the middle occupational group and at a greater distance than the Controls. If the dynamic

quality of the environment is in proportion to the intellectual level of the parents, then one would conclude that the Adoptive and Control homes are similar in stimulation potential. The difference, however, in intelligence of Adoptive children is only 1 IQ point in either direction, while the difference for the Control children is 5.8 points between occupational groups III and V and 10.7 points between groups III and II. The failure of Adopted children to attain levels of intelligence corresponding more exactly to those of the Control group would appear to be due to a factor or factors other than environment.

When the most stimulating environment was arbitrarily defined to exist in those homes that possessed all of the environmental traits at a level beyond 1 SD

TABLE 15

COMPARATIVE ANALYSIS OF ADOPTIVE AND CONTROL MID-PARENT SCORES ON THE OTIS TEST OF MENTAL ABILITY AND THE S.B. VOCABULARY TEST, CLASSIFIED ACCORDING TO OCCUPATIONAL STATUS

Occupational group	Adoptive parents Otis Test M	S.D.	S.B. Vocab. M	S.D.	Control parents Otis Test M	S.D.	S.B. Vocab. M	S.D.
I	59.6	8.0	74.0	6.4	64.6	5.4	74.9	7.8
II	59.6	6.7	73.4	7.2	57.1	10.0	67.8	8.3
III	49.6	11.9	64.6	11.4	51.8	11.5	62.0	9.3
IV	—	—	—	—	—	—	—	—
V	39.7	12.3	59.1	11.6	44.0	11.5	55.7	9.6
VI & VII	38.4	11.2	54.5	9.2	38.3	9.0	48.7	9.1

of the mean of the entire group, very interesting contrasts in mean IQ of the children of these homes appear. The seven Adopted children found in such homes had a mean IQ of 113.3±6.0. The eight Controls who were so located had a mean IQ of 127.5±9.5.

These Controls are 17.8 IQ increments above the mean (109.7) of their entire group; the Adopteds so classified are only 2.8 I.Q. points in advance of their mean (110.5). Despite the small number of cases involved, the ratio of the difference in IQ of these selected children to its standard error is 3.52. When the definition is reversed for the least stimulating environment, i.e., the homes located below minus 1 SD for every trait, no cases were found.

If we widen our definition of a stimulating environment so as to include the homes that were at the mean and above for every trait, 58 Adopted and 52 Control children are found. Here again the Control children surpass the Adopteds. The mean for the Controls is 119.4 ± 14.9, for the Adopteds 112.3 ± 10.8. The ratio of the mean difference to its standard error is 2.8. Hence, irrespective of attempts to equalize environments, the Controls from the upper levels are distinctly superior in intelligence to the Adopted children at the same levels.

When the children in the least stimulating environment (defining the latter to include the homes that score below the mean on every trait) are considered as a group the Adopteds secure a mean IQ of 106.0 ± 10.2, the Controls 99.5 ± 10.9. Although the numbers involved in this analysis are relatively small, 11 Adopteds and 16 Controls and the ratio of the mean difference to its standard error is only 1.5, the results are consistent with the comparisons made on the basis of occupation. Whether the impact of a poor environment is responsible for the lowered IQ is, of course, problema-

tic. Note that the Adopteds fall 4.5 IQ points below the mean of their entire group, while the Controls drop 10.2 points below the mean of their group. If environment is dominant one would expect the same amount of depression in IQ for both groups. It should also be noted that the difference between Control children of the upper and lower environmental levels as defined above is 19.9 IQ points. The difference in the case of the Adopteds is 6.3.

What is the explanation of the difference in IQ with variation in environment in these two sets of children? The marked similarity of the children in age, school grade, and mean IQ will be recalled. Similarity in other respects, also, exists. Vocabulary and IQ correlate .64 for Adopteds; 63 for Controls; Woodworth-Mathew scores of emotional stability and IQ correlate .02 for Adopteds, .06 for Controls. To what extent our results may be due to personality factors of which we have no measure it is impossible to say. We have no reason to believe that such attributes as self-confidence, industry, "drive" or their converse are differently distributed in the two groups of children. Nor have we any reason to believe that the emotional environment provided by the parents of the Control children is more stimulating than that provided by Adoptive parents. To the extent that coefficients of assortative mating (Table 16) are an index to qualitative factors in the home influencing the child's response to his environment there are no differences in our two sets of homes.

If selective placement is entirely absent in the Adopted population, variation in environment may be

TABLE 16

COEFFICIENTS FOR ASSORTATIVE MATING

	Adoptive parents			Control parents		
	r	P.E.	N	*r*	P.E.	N
Height	.26	.05	178	.27	.05	146
Education	.59	.03	192	.71	.02	193
Father's occupation and mother's education	.58	.03	192	.64	.03	194
Otis score	.57	.03	177	.41	.04	173
Vocabulary	.61	.03	174	.43	.04	164

said to be accountable for changes in IQ to the magnitude of about plus or minus 3 to 5 points. If, as previously stated, adoptive parents in the higher occupational, educational, and social levels secure children of greater promise than the adoptive parents in the lower levels, then the observed differences are clearly a function of genetic diversity rather than a function of environmental variation. Evidence relative to selective placement will be presented in the following section.

VII

Validity of Results

The confidence that can be placed in our results is dependent on: (1) the freedom of the data from constant errors of selection, (2) the soundness of our method, (3) the support available in supplementary evidence, (4) the absence of selective placement.

Selection

Inferences as to the existence of constant errors of selection may be made if on the analysis of the homes that refused to cooperate the presence of characteristics which would have definitely changed our results are discovered. Lower correlations between child's intelligence and environmental attributes would have been found if the refusals were for the most part from homes in which the child's mentality was widely separated from the cultural level of the home and conversely, higher correlations would result if the refusals were from homes in which the child's ability was in close agreement to the cultural level of the home. We have no reason to believe that either of these conditions prevailed. Information as to the school grade of 34 of the 37 refusals showed 29 at grade for age, 3 accelerated and 2 retarded. Further, judging from the occupational classification of the adoptive fathers, the children were fairly evenly distributed in all levels. In occupational group I the number of refusals was 6, 10 in II, 10 in III, 10 in V, 1 in VI and VII. Only 3 adoptive parents stated that they were unwilling to have

their child given a mental examination, and 2 of these parents declared their opposition to be a matter of religious conviction. Irrespective of the genesis of their excuse, the total number refusing on the basis of antipathy to a mental test is insignificant. Fear that the knowledge of adoption would reach the child and the fact that the child was sensitive about being adopted were the two most frequently given explanations for refusing cooperation. We have no evidence that either are related to the intelligence of the child or the cultural level of the home. Emotional upsets because of adoption seem probable when the age of the children in homes refusing cooperation is considered. On the average the children of these homes were one and a half years older (age 10.9±2.4) than the experimental children. In general, there seems little evidence that the coefficients of resemblance between adoptive parents and children were affected by our refusals.

Although, as will be recalled, the application of Pearson's formula corrected our curtailed variability and made the coefficients of resemblance for the Adopted children exactly comparable with those for the Control children, we are still confronted with the question of whether the variability in the entire universe (dependent adopted children) would be greater or less than that observed. Is our experimental population typical of the intellectual ability of adopted children? Is there anything in the data that would force the average level below or above that of dependent children in general? Comparing family backgrounds (Tables 17 and 18) from evidence of occupa-

TABLE 17

COMPARATIVE ANALYSIS OF THE OCCUPATIONAL DISTRIBUTION OF TRUE PARENTS OF ADOPTED CHILDREN

Occupational group	True fathers General group* N—1323	True fathers Experimental group N—89	Males in gen'l Pop. age 20-24 N—4, 121, 914	True mothers General group* N—1308	True mothers Experimental group N—89
I	2.1	1.1	1.7	0.4	0.0
II	4.5	2.2	2.4	4.0	9.0
III	18.1	27.0	14.1	8.7	15.7
IV	9.0	12.4	9.8		
V	35.8	23.6	32.0	25.8	31.5
VI & VII	30.5	33.7	40.0	61.1	43.8

*From Leahy (6, pp. 314, 324).

TABLE 18

COMPARATIVE ANALYSIS OF THE EDUCATIONAL ATTAINMENT OF TRUE MOTHERS OF ADOPTED CHILDREN

School attainment	General group N—1218*	True mothers Experimental group N—96
Completed high school or more	16.1	21.9
High school not completed	26.9	23.9
Less than high school	56.9	54.1

*From Leahy (6, pp. 307).

tion and education of parents with that of a larger sample of parents whose children were adopted leads one to the conclusion that the congenital level of the offspring of parents represented in our study would be similar to that of all adopted illegitimate children. It is possible that the true fathers and mothers whose occupation and education are unknown would either raise or depress the level of the group. However, the mean IQ of the children whose true father's occupation is known is 110.9, only .4 in advance of the mean of the entire experimental population. When the group

for whom either true parent's education or occupation is known are regarded separately, the mean IQ of the children is 110.8±12.5. Foundlings and children for whom the social agency had no family history, comprising a group of 48, have a mean IQ of 109.9±12.3. Hence, we may infer a similarity in occupational and educational level for the parents for whom we have no history.

The distortion of average IQ of adopted illegitimate children is most probably due to the fact of selected heredity. Illegitimate children who are not relinquished for adoption have in contrast an inferior social background (Leahy, 6). Further, as previously stated, laws which have been in force since 1925 in Minnesota do not permit the adoption of infants of parents known to be feebleminded. Since 1918, a trial period of six months' residence in the new home before legal adoption has been mandatory. Such policies doubtless raise the average level and reduce the range of intelligence of children who are placed in adoptive homes. It seems fair to conclude that our experimental children are no more or less variable in intelligence than adopted illegitimate children in general. The fact of innate positive skewness in intelligence should not prevent or diminish the operation of the influence of environment on intelligence. If environment is prepotent we should find significant variability in intelligence with variability in environment.

Method

Only insofar as our method can be demonstrated to

have avoided ambiguities, to have insured accuracy, and to have provided a check upon the reliability of our measures may we have confidence in our results. The first of these was fulfilled by the criteria held for experimental subjects. You will recall that every precaution was taken to escape ambiguities arising from selective mental resemblance between adoptive parent and child. Age at placement, race, and nationality extraction were rigidly controlled.

In order to insure accuracy uniform procedures were adhered to in gathering the social data, in the administration of tests, in scoring tests, and in analyzing the data. As previously reported two experienced examiners whose training and experience were practically identical administered all of the children's tests. All tests were rescored. All classifications and statistical computations of a major nature were done on the Hollerith Sorting and Tabulating Machine. Computations involving the use of calculating machines were systematically checked. A check upon the reliability of our measures, as well as on the whole experimental procedure, was provided by the Control group. Since these children agreed exactly in number, age, and sex with the experimental subjects, we have reason to believe that our tests and measures reflect the true ability of both groups with equal accuracy. The extent to which Control homes were truly a control is attested by the marked agreement of the two sets of homes (adopted and control) for each environmental attribute that we attempted to measure. Hence, we may conclude that the results from the Adoptive and Control groups furnish a valid comparison.

SUPPLEMENTARY DATA

If environment is dominant we should expect that unrelated children in the same household would agree markedly in ability. The contrary was found as may be seen by the following:

	r	PE	N
IQ of Exper. Adopteds and Own Children	.06	.14	25
IQ of Unrelated Adopteds in Same Household	.12	.22	10
Vocab. of Exper. Adopteds and Own Children	.06	.13	25

Although the number of cases is small the results suggest that the children are widely different in intellectual ability, regardless of their common environment.

In 20 cases of own children of Adoptive parents, IQ of own child and mid-parent Otis correlated .36. This correlation follows the expected familial pattern.

Although our single measure of the emotional stability of the children is probably not sufficiently reliable to permit any conclusions about personality differences, the similarity in the relationship between Woodworth-Mathews' scores and home environment for both groups of children is striking. As indicated below, Woodworth-Mathews' scores correlate:

.10 for Adopteds, .13 for Controls with mid-parent Otis Score

—.04 for Adopteds, .07 for Controls with mid-parent Vocabulary

.06 for Adopteds, .13 for Controls with Cultural index of home

.11 for Adopteds, .18 for Controls with Child Training Index of home

—.04 for Adopteds, .07 for Controls with Occupation of father.

Clearly the fact of either blood relationship to persons shaping the environment or its absence makes no difference. Adopted children, where presumably only environment is operative, behave in a manner similar to own children where both heredity and environment are operative. These results are in distinct contrast to our observations on the relationship of intelligence to home environment. They are based on the tests of 72 Adopteds and 77 Controls and give support to the theory that heredity plays a less significant role than environment in the variation observed for traits other than intelligence

TABLE 19

WOODWORTH-MATHEWS SCORES ACCORDING TO OCCUPATIONAL STATUS

Occupation of father	Adopted children N	M	S.D.	Control children N	M	S.D.
I	16	59.3	6.8	11	63.1	4.0
II	9	54.8	9.7	14	61.1	6.0
III	17	59.7	7.5	19	61.1	5.2
IV	—	—	—	—	—	—
V	23	58.8	7.9	25	62.5	5.7
VI & VII	7	60.4	6.7	8	59.3	8.6

Table 19 demonstrates the similarity in performance on the Woodworth-Mathews when the children are classified according to the occupation of their fathers. The cross differences are slight; none are significant. Further, it should be noted that there is no significant progression in score with occupational status in either

group. This is directly contrary to our observations for test intelligence of children and occupational status of father. As may be seen by referring to Table 12, although Adopted children did not progress significantly in IQ with occupational status, the Control children made significantly large gains in IQ when comparisons were made between the two lowest and middle occupational groups and when the latter was contrasted with the two highest levels.

Selective Placement

The extent to which our results are free from selective placement of children on the basis of cultural similarity between adoptive and true parents determines the limits of the influence of environment on intelligence. If selective placement is entirely absent, then our results stand as evidence of the maximum influence of environment. If selective placement exists then our results overstate the influence of environment on intelligence.

In Table 20 the relationships of cultural background of true and adoptive parents are shown. All of the coefficients are low. Only three are reliable. The majority, however, are positive. In general, they offer weak evidence for selective placement in our data. Nevertheless, further analysis was applied for an appraisal of the point.

One check involved an analysis of the replies of 22 placement workers to a questionnaire which asked the importance of certain attributes in judging the fitness of a child for a home. Most of the workers who an-

TABLE 20

FACTORS IN ADOPTIVE HOME CORRELATED WITH FACTORS IN THE BACKGROUND OF CHILD AS REPORTED BY SOCIAL AGENCY

	True father's occupation			True mother's occupation			True mother's education		
	r	P.E.	N	r	P.E.	N	r	P.E.	N
Adoptive father's occupation	.09	.07	89	.20	.07	89	.23	.06	96
Adoptive father's education	.24	.07	88	.21	.07	88	.20	.07	95
Adoptive mother's education	.10	.07	88	.38	.06	88	.25	.07	94
Adoptive mid-parent Otis score	.11	.07	79	.12	.07	83	.20	.07	89
Adoptive mid-parent vocabulary test	—.02	.08	76	.07	.07	79	.08	.07	86
Social status of adoptive home	.20	.07	89	.29	.06	89	.23	.06	96

swered the questionnaire had been engaged in child placing in Minnesota during the period from which our cases were drawn. Eighteen of the 22 regarded *probable intelligence of child* as of very great importance. The group, however, was evenly divided on the significance of education and occupation of true parents in judging *probable intelligence of child*. Eleven regarded education of true mother of slight or moderate importance and 11 held it to be of great or very great importance. The ratings for education of true father followed the same pattern. The judgments on the significance of occupation of true parents were similarly divided. Obviously if these placement workers behave in accordance with their replies the possibility of selective placement in a group of adopted children on the basis of education and occupation of true parents is nil.

The second check consisted in holding background constant in the correlation between child's IQ and adoptive home. This, clearly, is the crucial test. Time

permitted the analysis for only a limited number of traits of the adoptive home. The traits selected, however, and presented in Table 21, are the most generally

TABLE 21
COMPARATIVE CORRELATIONS BETWEEN CHILD'S IQ AND ADOPTIVE HOME

	Background factor held constant		Zero order r background not constant	
	True father's occupation N=89	True mother's occupation N=89	True mother's education N=96	Entire group
Adoptive father's occupation	.18	.03	.05	.12
Adoptive father's education	.12	.13	—.02	.16
Adoptive mother's education	.18	.16	.13	.21
Adoptive mid-parent Otis score	.26	.10	.09	.18
Adoptive home environment score	.19	.16	.02	.19

used indices of environment; namely, occupation, education, test intelligence of parents, and home rating.

An examination of this table shows that the correlations between child's IQ and attributes in the adoptive home are not greatly different from those for the entire group except when true mother's education is held constant. Then the correlations drop for each correlated factor. What may we infer? Clearly, for the populations involved, true mother's education appears to be a basis of selective resemblance between the adoptive home and its child, while true parent's occupation does not. Whether the population for whom no background history is available would support or refute these findings is problematic. Certainly the correlations reported in Table 11 may be regarded as no understatement of the relationship between

child's IQ and home environment. From evidence of family background reported by adoptive parents on 50 of our cases, mental resemblance due to selective cultural likeness seems highly probable and therefore our correlations for child's IQ and environment tend to overstate the influence of nurture. Because of the possibility of retrospective falsification in reports of this kind, these data may be ignored and the correlation of .20 be regarded as a general characterization of the relationship of child's IQ and environment.

VIII

CONCLUSIONS

By methods which allowed the effects of environment to be studied separately from those of heredity in combination with environment, this study attempted to discover the influence of environment and heredity on intellectual variation. As stated in the opening section, the tendencies observed in this study are valid only for populations which are similar to the experimental population in composition. However, the consistency with which a coefficient of .50 was secured for parent and offspring suggests that the restricted range in both the hereditary and environmental variables was reciprocal and hence no serious distortion in our results exists. The main conclusions are as follows:

1. Variation in IQ is accounted for by variation in home environment to the extent of not more than 4 per cent; 96 per cent of the variation is accounted for by other factors.

2. Measurable environment does not shift the IQ by more than 3 to 5 points above or below the value it would have had under normal environmental conditions.

3. The nature or hereditary component in intelligence causes greater variation than does environment. When nature and nurture are operative, shifts in IQ as great as 20 IQ points are observed with shifts in the cultural level of the home and neighborhood.

4. Variation in the personality traits measured in this study other than that of intelligence appears to be accounted for less by variation in heredity than by variation in environment.

References

1. Burks, B. S. The relative influence of nature and nurture upon mental development. A comparative study of foster parent-foster child resemblance and true parent-true child resemblance. *27th Yrbk. Nat. Soc. Stud. Educ.*, 1928, Pt. I, 219-316.

2. Freeman, F. N., & Others. The influence of environment on the intelligence, school achievement, and conduct of foster children. *27th Yrbk. Nat. Soc. Stud. Educ.*, 1928, Pt. I, 103-217.

3. Goodenough, F. L., & Anderson, J. E. Experimental child study. New York: Century, 1931. Pp. xii+546. (See esp. 501-512).

4. Kelley, T. L. Statistical method. New York: Macmillan 1923. Pp. xi+390.

5. Lawrence, E. M. An investigation into the relation between intelligence and inheritance. *Brit. J. Psychol., Monog. Suppl.*, 1931, 16, 1-80.

6. Leahy, A. M. A study of certain selective factors influencing prediction of the mental status of adopted children. *J. Genet. Psychol.*, 1932, 41, 294-329.

7. Schwesinger, G. C. Heredity and environment. New York: Macmillan, 1933. Pp. v+484.

8. Terman, L. M., & Others. Genetic studies of genius, Vol. I. Mental and physical traits of a thousand gifted children. Stanford University, Calif.: Stanford Univ. Press, 1925. Pp. xv+648.

LA NATURE—LE MILIEU ET L'INTELLIGENCE

(Résumé)

Au moyen d'une comparaison d'un groupe d'enfants adoptés avec de vrais enfants vivant dans un milieu approximativement pareil, qui a permis que les effets du milieu soient étudiés séparément de ceux de l'hérédité en combinaison avec le milieu, cette étude a essayé de découvrir l'influence du milieu et de l'hérédité sur la variation intellectuelle. Les tendances observées dans cette étude sont de valeur seulement pour les populations qui sont d'une composition pareille à celle de la population expérimentale. Cependant, la constance avec laquelle on a obtenu un coefficient de 0,50 pour le parent et l'enfant suggère que la variation restreinte dans les variables héréditaires et celles du milieu a été réciproque et il n'existe ainsi nulle distorsion sérieuse dans nos résultats. Les conclusions principales sont les suivantes:

1. La variation dans le QI est expliquée par la variation dans le milieu de la maison de non plus de quatre pour cent; 96 pour cent de la variation est expliqué par d'autres facteurs.

2. Le milieu mesurable ne change le QI que de 3 à 4 points au-dessus ou au-dessous de la valeur qu'il aurait eue dans les conditions normales du milieu.

3. La composante de nature ou héréditaire dans l'intelligence cause une plus grande variation que celle causée par le milieu. Quand la nature et le milieu sont opératifs, des changements du QI aussi grands que 20 points du QI sont observés avec des changements dans le niveau de culture de la maison et du voisinage.

4. La variation dans les traits de personnalité mesurés dans cette étude autres que celle de l'intelligence semble être expliquée moins par la variation dans l'hérédité que par la variation dans le milieu.

LEAHY.

VERERBUNG, UMGEBUNG UND INTELLIGENZ

(Referat)

Durch einen Vergleich einer Gruppe von adoptierten Kindern mit anderen Kindern in ungefähr derselben Umgebung, der die Untersuchung der Wirkungen der Umgebung gesondert von denjenigen der Vererbung in Verbindung mit der Umgebung ermöglichte, wurde in dieser Untersuchung der Versuch gemacht, den Einfluss der Umgebung and Vererbung auf intellektuelle Abweichung aufzudecken. Die Ergebnisse dieser Untersuchung sind gültig nur für Gruppen, welche der experimentellen Gruppe ähnlich sind. Die Gleichförmigkeit aber, mit der ein Koeffizient von 0,50 für Eltern und Nachkommenschaft erhalten wurde, weist darauf hin, dass der beschränkte Umfang bei den Umgebungs- und Vererbungsvariablen wechselseitig waren und folglich keine Verstellung unserer Ergebnisse aufweisen. Die Hauptschlüsse sind folgende:

1. Eine Variation des IQ wird durch die Variation der Heimumgebung bis zum Umfang von mehr als 4 Prozent erklärt; 96 Prozent der Variation wird durch andere Faktoren erklärt.

2. Die messbare Umgebung verändert den IQ nicht mehr als 3 bis 5 Punkte über oder unter dem Wert, den er sonst unter normalen Umgebungsumständen haben würde.

3. Die Natur- oder Vererbungskomponente bei der Intelligenz verursacht eine grössere Variation als die Umgebung. Wenn Vererbung und Umgebung tätig sind, werden Verschiebungen des IQ so gross wie 20 IQ-Punkte mit Verschiebungen der Kulturschichten des Heims und der Nachbarschaft beobachtet.

4. Die Variation des Persönlichkeitszuges ausser der Intelligenz, der in dieser Untersuchung gemessen wurde, scheint weniger durch die Variation der Vererbung als durch die Variation der Umgebung erklärt zu werden.

LEAHY.

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

October, 1935
Volume XVII, No. 5

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

October, 1935
Volume XVII, No. 5

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

ON INTELLIGENCE OF EPILEPTIC CHILDREN*

From the Department of Psychology of the University of California at Los Angeles

By

ELLEN B. SULLIVAN AND LAWRENCE GAHAGAN

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, February 8, 1935.

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

CONTENTS

I

INTRODUCTION

This report is concerned primarily with the intelligence and secondarily with the personality and conduct of an unselected group of 103 epileptic children admitted to the Children's Hospital, Los Angeles, California, and referred to the Psychological Clinic for study.[1]

The Department of Psychology of the University of California at Los Angeles has conducted a psychological service clinic at the Children's Hospital since 1925. During that time examinations have been made of more than 900 cases referred for various problems and suffering from a variety of physical diseases or disabilities. The psychological examination has been a part of the medical service, or in cases referred for psychological and educational service, has been supplemented by medical examination. Results having research value have accumulated and this and other reports to follow will present the study of accumulated data.

---

[1]This study was made possible by the friendly cooperation of the hospital staff, and particularly the following persons: Doctors Howard Cooder, Barnett Lipson, J. M. Nielsen, Samuel Ingham, and Helen Hopkins.

II

DESCRIPTION OF CASES STUDIED

Age, Sex, Race, and Physical Facts

The group suffering from epilepsy included all the cases of this type examined by us. The Hospital does not accept cases of children over twelve years of age except when follow-up of their own cases is required. The study is therefore concerned predominantly with a group of younger children. Tables 1, 2, and 3, which follow, give the age, sex, and race of the epileptic group.

TABLE 1

Sex of Cases Studied

Males	63
Females	40
Total	103

TABLE 2

Race of Cases Studied

White	American Non-Jewish	72
	American Jewish	7
	Mexican	19
	Armenian	1
	Italian	1
Japanese		2
Negro		1
	Total	103

TABLE 3
CHRONOLOGICAL AGES OF CASES STUDIED*

Age in years	Number of cases
14	1
13	1
12	8
11	11
10	10
9	12
8	6
7	13
6	10
5	5
4	16
3	5
2**	5
Total	103

Mean of group:	7 years 6½ months
Median:	7 years 4 months
Q_3	10 years 0 months
Q_1	4 years 6 months
Range	1 year 10 months to 14 years 9 months

*Ages reported are computed at time of first psychological examination.
**Two years = interval 1 year 6 months to 2 years 5 months, etc.

We have investigated the distribution of intelligence of these children and have sought to discover what relationships, if any, exist between intelligence and such factors in epilepsy as causation, severity, age at onset, duration, and cessation of attacks. We have also sought to discover whether or not there exist modes of behavior which might be described as typical of epileptic children.

Epilepsy[2] may be defined as "the sudden and re-

[2]For a general discussion of epilepsy, see Lennox and Cobb (17, pp. 105-290). These authors state that "Little is really known of the neurological mechanism involved in an epileptic seizure" (p. 129). They state four theories of the neurological mechanisms of convulsions: (1) the "irritation" theory, (2) the "release" theory,

peated appearance of seizures of which convulsive movements or loss of consciousness, or both, are the principal elements" (Cobb, 3, p. 1245). It is a phenomenon of varying causation as observed clinically. Cobb (3, pp. 1245-1263), for example, gives 56 clinical causes of convulsions, and postulates 13 physiological mechanisms to explain their occurrence. The incidence of epilepsy in the whole population has never been adequately established. Malzberg (18, p. 97 f.), after reviewing various statistical sources, states that approximately five epileptics per 1000 general population is a fair estimate. Likewise the sex ratio of epilepsy is undetermined.[3] The greater frequency of male to female cases in institutions may reflect not the actual incidence but the extent to which institutional provision has been required (30, p. 658). Our data bears inconclusively upon the question of sex ratio of epilepsy. The 103 cases, with which this report is concerned, are divided into 63 boys and 40 girls. This shows a decided numerical superiority of males. In order to pursue this point it is necessary to know the sex ratio of all cases admitted as out-patients or in-patients to Children's Hospital for the approximately corresponding period of years. This comparison is given in Table 4 which shows a larger proportion of males

---

(3) the "short-circuit" theory, and (4) the "explosive" theory (pp. 111-130). Cobb (3, p. 1247) suggests the existence of some constitutional factor in persons who develop epilepsy. Penfield and Gage (29, pp. 709-727) have described the cerebral localizations of epileptic manifestations in 75 cases of focal epilepsy.

[3]For a discussion of the sex ratio in epilepsy, see Pollock (30, p. 658).

TABLE 4
SEX RELATIONSHIP

	Percentage males	Percentage females
Total hospital population (new cases) 31,416 cases, 1930-34	54.2	45.8
Epileptic group 103 cases	61.2	38.8

in the epileptic group than in the general population of the new cases admitted to the Children's Hospital.

CAUSATION OF EPILEPSY IN CASES STUDIED

With reference to causation, epilepsy is divided into organic (or symptomatic) and idiopathic (or essential) epilepsy. Table 5 summarizes the data on causation in our cases.[4]

TABLE 5
CAUSATION

Cause	Number of cases
Organic	
(*a*) Definite	17
(*b*) Questionable	15
Idiopathic*	
(*a*) Undifferentiated	53
(*b*) Psychogenic	2
No cause given	16
Total	103

*One case included in our data as idiopathic subsequently developed focal symptoms. At operation a deep left parietal tumor was found.

[4]Comparison of causation of epilepsy of our group with an institutional group (Craig Colony, New York) follows. Of the 356 first admissions to the Craig Colony for the year ending June 30, 1933, 39.8 per cent were classified as symptomatic or organic, and 60.2 per cent as idiopathic. The corresponding percentages in the previous year were 47.7 per cent and 52.3 per cent, respectively (23, p. 335).

Table 6 subclassifies the cases with organic etiology (definite and questionable).

TABLE 6

CLASSIFICATION OF ORGANIC CAUSES*

	Number
Definite	
Jacksonian**	5
Cerebral birth injury	4
Post natal cerebral injury	3
Encephalitis	2
Encephalitis and pneumonia	1
Congenital syphilis	1
Organic (without description)	1
Questionable	
Cerebral birth injury	7
Post natal cerebral injury	2
Encephalitis	6
Hydrocephalus	2
Fever	1
Organic (without description)	1

*Causation is recorded from the diagnostic statement of the neurologist examining case. The apparent discrepancy in the total number of the questionably organic cases in Tables 5 and 6 is due to the listing of multiple causation in Table 6.

**Sargent (32, p. 321) defines a Jacksonian convulsion as "Merely a series of fractional movements, devoid of purpose, and determined in point of sequence only by the relative anatomical position of the cortical areas in which they are represented." Of 270 cases of brain tumor, exclusive of cerebellar and pituitary cases, upon which Sargent operated, 30 per cent exhibited Jacksonian convulsions (32, p. 326).

Two of our idiopathic cases (Table 5) are described as psychogenic. Comparing our cases with institutional cases (Craig Colony), for the year ended July 1, 1925, of 225 admissions, five reported a psychogenic etiology (20, p. 10), and for the year ended June 30, 1933, of 356 first admissions, two reported psychogenic factors (23, p. 335). The small incidence of psychogenic epilepsy in our group is in agreement with the findings of Craig Colony. Cookson (4, p. 185) says

that "the psychogenic conception will probably only be the explanation in a small number of cases, and would seem to have its most limited application in children, in whom the 'adjustment demand' is not likely to be so intolerable as to produce convulsions."

One of our definitely organic cases (Table 6) had its causation in congenital syphilis. The low incidence of syphilis in the production of epilepsy has been noted by Patterson (27, p. 66) who reported that of 2240 blood Wasserman tests done on epileptics at the Craig Colony Laboratory, only 40 (or about 1.8 per cent) gave a plus four reaction. Patterson concludes that the occurrence of syphilis among epileptics does not seem to exceed that in the general population. Syphilis, then, appears to be a relatively infrequent cause of epilepsy and of mental deficiency, whereas it is a frequent cause of insanity.

Seven of our cases (Table 6) received a definite diagnosis of cerebral injury, of which four occurred at birth and three post-natally. In addition nine other cases received a questionable diagnosis of cerebral injury, seven occurring at birth and two post-natally. Cerebral injury, natal and post-natal, plays an important rôle in the production of epilepsy as well as in mental deficiency. Gowers (14, p. 123) as early as 1881 cited cerebral birth injury as a cause of epilepsy. Twenty-eight cases of a total of 225 admitted to the Craig Colony during the year ended July 1, 1925, were given an etiological classification of head trauma, of which seven cases occurred at birth (20, p. 10). Shanahan (33, p. 38) reported that of 1500 admissions of

all ages to Craig Colony, 8 per cent gave a history of trauma at birth or shortly after. Rosanoff, Handy, and Rosanoff (31, p. 1172) state that "the relationship of epilepsy existing from birth or from an early age, to mental deficiency and to Little's disease suggests a common etiological factor for the three conditions, namely, trauma at birth." Myerson (1, p. 36) says that "even though we have reached no definite conclusion as to the mechanism of epilepsy, it is definitely realized that brain injury plays a larger rôle than we thought."

The significance of post-natal head injury in the production of epilepsy is well indicated in the report of Sargent (32, pp. 322-323), who states that of 18,000 individuals who had suffered gunshot wound in the head and who were examined by Re-survey boards of the Ministry of Pensions, nearly 4.5 per cent were subject to epileptic convulsions.[5] "Seeing, however, that only a very small number of patients with anchored brain develop epilepsy, it is obvious that the local lesion is but one of the links in the etiological chain"

[5]Stevenson (34, pp. 214-224) reviews the literature on the relation of post-natal head injury and epilepsy. The frequency of epilepsy following head injury as reported by Sargent, we may regard as conservative. This is in view of the fact that the re-survey examinations were made in 1919-1920, which would allow a maximal interval of approximately six years (i.e., from 1914) from the time of injury to that of onset of convulsions. That longer intervals may intervene between injury and onset is well known. Thus, Foerster and Penfield (11, p. 118) reported that of 12 cases of traumatic epilepsy, the time between the wound and the onset of convulsion varied between five months and fourteen years, with the average interval being 5 years and 6 months. Stevenson (34, p. 215) reported that the longest interval in his series of cases was 14 years, and in 11 cases, the interval was at least 10 years.

(32, p. 323). Drayton (7, pp. 1302-1309) has pointed out that personality changes, as well as headache and dizziness, are a common sequela of post-natal cerebral injury. The personality changes observed are "extreme irritability, inability to endure noise without emotional upset, marked restlessness, hyperactivity, emotional instability and distinctly antisocial behavior" (7, p. 1302). Encephalography revealed that the syndrome of headache, dizziness, and personality change is of organic rather than functional origin. Forty-three cases of post-traumatic personality disorder, the majority of whom were court cases, and some of whom were convulsive, were treated by lumbar insufflation of air. Of these, 49 per cent subsequently became normally adjusted, 37 per cent improved, and 14 per cent remained unimproved.

Rosanoff, Handy, and Rosanoff (31, p. 1172) state that in the majority of cases of epilepsy (as well as mental deficiency and Little's disease) no injury at birth is reported. A consideration of the birth histories of our cases shows not only birth injury but also the frequent occurrence of adverse birth factors, such as instrumental delivery, prematurity, etc. Table 7 and Table 8 summarize the birth histories of our cases. The data contained in these tables are taken from the birth histories as given in the hospital records as well as from reports of examining neurologists. One adverse factor, birth order, is not reported in our tables because the information with respect to it was not completely available.

TABLE 7

BIRTH HISTORY

	Number of cases
Normal	55
Adverse history (no injury proved)	41
Birth injury	5*
No birth history available	2
Total	103

*One case, in which the history reported birth injury, did not have a neurological examination and, therefore, Table 6 shows one less case (4) of birth injury than Table 7 (5).

All data in Tables 7 and 8 were obtained from birth history available as well as from neurological examination.

TABLE 8

NATURE OF ADVERSE BIRTH FACTORS

	Number of cases
Birth injury	5
Adverse history factors (total—41 cases)	
Instrument	12
Premature	8
Long labor	8
Immature	6
Difficult labor	5
Precipitate birth	4
Induced labor	3
Cesarian	3
Overweight	3
Breech delivery	2
Stuperous	1
Dry birth	1
Difficult presentation	1
Blue	1
Twin birth	1*
Difficult resuscitation	1

*Other twin died in convulsion in infancy.

NEUROLOGICAL PATHOLOGY

Table 9 refers to the neurological examination.[6] A large number and variety of pathological conditions were found, the condition occurring most frequently, namely Babinski-type reflexes, indicating lesion of the pyramidal tracts (16 cases). Other pathological conditions found in more than one case were: deep reflex increased (10 cases), choreo-athetosis (five cases), spasticity (five cases), hypotonia (three cases), and the following in two cases each: optic disc atrophy, pupillary signs, nystagmus, staggering gait, muscular weakness, and endocrinopathy (one case hypothyroidism, one case hypoparathyroidism). The high incidence of pyramidal tract lesions is in accordance with the rôle of cerebral injury and infection in the production of epilepsy.

TABLE 9
NEUROLOGICAL EXAMINATIONS*

	Number of cases
No neurological examinations	22
Neurological examinations, negative	37
Neurological examination, pathological findings	44
Total	103

*Neurological examination did not include encephalography nor ventriculography, which, had they been done, would have undoubtedly revealed more neuropathology.

SEVERITY OF ATTACKS

Classification of our cases in terms of severity of attacks is given in Table 10. The more severe or ma-

[6] Lennox and Cobb (17, p. 137 f.) review the results of neurological examinations of epileptics.

TABLE 10

SEVERITY OF ATTACKS*

Type	No. of cases
Major	54
Major and minor	17
Major, minor, and equivalent	2
Major and equivalent	3
Major, later minor	7
Total major (83)	
Minor	16
Minor and equivalent	3
Total minor (19)	
Hystero	1
Total hystero (1)	
Total	103

*Data concerning frequencies of attacks are omitted because they were not available in a sufficiently complete form to be statistically treated.

jor condition existed in over 80 per cent of the cases, which indicates a greater proportion of the major condition than usually found in a non-institutional series.[7]

TABLE 11

RELATION OF SEX TO SEVERITY*

	Number in total group	Major percentage	Minor percentage	Hystero percentage
Males	63	85.7	14.2	0
Females	40	72.5	25.0	2.5

*Showing a greater proportion of the major condition in males. Also showing the presence of one case of hystero-epilepsy in females.

[7]In an institution group, the proportion of patients with major seizures, either alone or in conjunction with minor seizures, is relatively high. Thus of new cases admitted to the Craig Colony in the two years ended June 30, 1933, over 94 per cent were subject to major attacks (23, p. 336).

Cookson's (4, p. 180) series of 100 epileptic children, for example, contain about equal numbers of major and minor cases. Table 11 classifies our cases in relationship of sex to severity, showing a greater proportion of the major condition in boys than in girls. This difference appears to be significant. Tables 10 and 11 show one case, that of a girl, diagnosed as hystero-epilepsy.[8]

Age at Onset

Tables 12, 13, and 14 deal with the age at onset of convulsions in our series of cases.[9] Table 12 shows the range of ages at onset and the number of cases at each age level. Table 13 expresses the same data in terms of the cumulation of occurrences at successive levels. Table 14 shows the cumulation of occurrences in the sexes listed separately.

---

[8]Worster-Drought (38, pp. 50-82), who gives a general discussion of the relationship of hysteria and epilepsy, says, "It is by no means infrequent to meet with an almost pure form of both major epileptic seizure and hysterical convulsion in the same individual."

[9]Of 6075 cases of known age at onset admitted to the Craig Colony, 30 per cent had their onsets from birth to five years of age, 18 per cent had their onsets from five to 10 years of age, and 19 per cent had their onsets from 10 to 15 years of age (20, p. 14). Thus, 67 per cent of institutional cases had onset in infancy and childhood. Shanahan (33, p. 38) reports that of 7000 institutional admissions, the age at onset in 28 per cent of the cases was before five years. Paskind (26, p. 382 f.), citing several studies of the age of onset of epilepsy, says that his extramural group of 304 cases contained a smaller percentage of patients with onset before the age of 5 years, (namely, 11.6 per cent) than any other series.

TABLE 12

AGE AT ONSET*

SHOWING THE RANGE OF AGE AND THE NUMBER OF CASES AT THE GIVEN AGE

Age range	Number of cases having onset at given age
11:6-12:5	2
10:6-11:6	3
9:6-10:6	5
8:6- 9:6	2
7:6- 8:6	5
6:6- 7:6	8
5:6- 6:6	8
4:6- 5:6	5
3:6- 4:6	8
2:6- 3:6	12
1:6- 2:6	11
0:6- 1:6	22
0- 0:6	11
Total	102

*This table shows a tendency toward early onset. Age at onset undetermined in one case.

TABLE 13

SHOWING CUMULATION OF AGE AT ONSET AT SUCCESSIVE AGES*

Age of onset	Number of cases	Percentage of cases
12 years or sooner	102	100
11 years or sooner	98	96.0
10 years or sooner	96	94.1
9 years or sooner	91	89.2
8 years or sooner	87	85.2
7 years or sooner	82	80.3
6 years or sooner	75	73.5
5 years or sooner	69	67.6
4 years or sooner	63	61.7
3 years or sooner	53	51.9
2 years or sooner	42	41.1
1 year or sooner	25	24.5
6 months or sooner	14	13.7
1 month or sooner	4	3.9
Total 102 cases (62 Males, 40 Females)		

*The table shows that approximately 25 per cent of our cases had onset of convulsions at one year or sooner, over 50 per cent at three years or sooner, and over 80 per cent at seven years or sooner.

TABLE 14

AGE OF ONSET AND SEX*

Age before which a given percentage of the cases occur	Percentage occurring before a given age Males	Females
12 years 6 months	100	100
11 years 6 months	96.7	100
10 years 6 months	93.5	90.75
9 years 6 months	88.7	90.5
8 years 6 months	87.0	81.0
7 years 6 months	82.2	80.5
6 years 6 months	75.7	75.0
5 years 6 months	67.7	60.7
4 years 6 months	66.1	57.5
3 years 6 months	58.0	50.0
2 years 6 months	45.1	40.0
1 year 6 months	33.8	30.0
6 months	11.2	10.0
1 month	6.4	0.
Total cases included (62 males, 40 females)		

*This table shows a slightly greater percentage of early onset in males than in females. This difference may not be significant.

III

RESULTS OF OUR STUDY OF INTELLIGENCE AND PERSONALITY OF EPILEPTIC CHILDREN

Intelligence of Group Studied

The problems related to the intelligence of epileptic children are numerous and of great practical and theoretical significance. It is important to arrive at some reliable picture of the group as a whole as well as to discover the variability within the group.

A number of studies have been reported which tend to indicate that the group as a whole is significantly retarded mentally. Gowers (14, p. 122 f.), in 1881, found that 64 cases of a total series of 1450 presented "considerable mental defect." Fox (12, pp. 235-248) investigated, by means of the Binet-Simon (Burt's revision) and other intelligence and educational tests, the intelligence of a group of 150 epileptic children (99 boys and 51 girls), who were resident in an English epileptic training colony. These cases were somewhat selected, the applicants for admission presumably being "capable of some education and occupation, and not defective within the meaning of the Mental Deficiency Act" (12, p. 235). The children's ages ranged from five to 16 years. Fox found the Median Binet Intelligence Quotient to be 71 for the boys, and 65 for the girls. His results lose some value because of the selected nature of his subjects. He states also (p. 240 f.) that "tests which epileptics find relatively

harder (than the normal child) include those where immediate memory is concerned, where written language is especially involved, where there are abstract questions to be solved, calling for concentration and reasoning, and finally where the use of coins as in every day is tested."[10] He further states that in general the failures of Binet test items among epileptic children were similar to those among mentally deficient children.

Ninde (24, pp. 1-36) tested the memory span for digits of a group of 2000 institutional white epileptics of age range from five to 50 years, and found the memory span to be approximately one-half that of normal individuals of corresponding ages. The relative ability of one-half normal was found for both the child and the adult, indicating an apparent absence of progressive deterioration of these subjects. Cookson (4, p. 178) states that 26 cases from a group of 100 epileptic children in the Royal Hospital for Sick Children, Glasgow, showed a "deficiency of mentality, more or less marked; in all the others it was normal. When mental deficiency was present, it had been noticed as early as the fits." He does not describe his method of measuring mental deficiency.

Dawson and Conn (5, pp. 142-151) later, using Burt's revision of the Binet-Simon test, measured the intelligence of 49 epileptic children in the same hospital. The ages ranged from four to 12 years. Twenty-one cases were retested after intervals of eight

[10]The deficiency in the use of coins is characteristic of institutional children, who lack opportunity to use money.

months to four years eight months. The Mean Intelligence Quotient of the group on the first test was found to be 80.65, the P.E. of the Mean being 1.43. The range of Intelligence Quotients was 49 to 117, the frequencies for the various intervals as follows: 110-117, three cases; 100-109, three cases; 90-99, seven cases; 80-89, 12 cases; 70-79, 13 cases; 60-69, four cases; 50-59, six cases; 40-49, one case. He states that the Mean IQ of the epileptics was "appreciably and significantly below the average of the rest of the hospital children examined" (991), which is 90.57 (P.E. mean $\pm$.31). Excluding cases of patients suffering from diseases of the ductless glands and of the brain, the Mean Intelligence Quotient of the rest of the patients is 91.76 (P.E. Mean $\pm$.35) which is still more above that of the epileptics. The standard deviation of Intelligence Quotients of the group of epileptics (14.87) was similar to that of other cases (14.66). The difference did not prove to be statistically significant.

The annual reports of the Department of Mental Hygiene of the State of New York have for several years summarized the intelligence status of epileptic persons newly admitted to Craig Colony. For the year ended June 30, 1931, of 317 first admissions, only 21.4 per cent were classified as of normal mental status, and 78.2 per cent were classified as of some degree of mental subnormality (21, p. 336). Of the subnormal group, 10.7 per cent were classified as idiot, 21.4 per cent as imbecile, and 44.5 per cent as moron. A marked sex difference was noted, only 12.6 per cent of the females being classified as normal as compared with 28.7

per cent of the males. The method of mental measurement is not given in the New York reports. Relatively the same results are given for the two following years. For the year ending June 1932, reports show 24.2 per cent normal and 75.9 per cent subnormal (22, p. 339). Of the subnormal, 4.6 per cent were idiots, 24.2 per cent imbeciles, and 47.1 per cent morons. For the year ending June 1933, reports show (23, p. 338) 24.4 per cent were normal. Of the subnormal group, 6.7 per cent were idiots, 14.6 per cent imbeciles, and 53.9 per cent morons. The same sex differences continued showing a larger percentage of females mentally retarded than males.

The high percentage of subnormal patients in the institutional epileptics as compared with extramural cases indicates the selected nature of the institutional group. Paskind (26, pp. 370-385), for example, reports that of 304 cases from private practice, who had been epileptic for at least six years, only 6.5 per cent were found to be mentally deteriorated. His method of determination of deterioration is not given. He says "The remainder or 93.4 per cent were found to be in excellent mental health, without the slightest trace of deterioration or other psychosis, and engaged in occupations similar to those of the great mass of the population" (p. 373).

Fetterman and Barnes (10, pp. 797-801), using the Stanford-Binet test found the average Intelligence Quotient of 105 epileptic patients (ages not given) of a hospital dispensary group to be 74, with a range from 34 to 133. Forty per cent were found to be mentally

defective (below 70 IQ), 18 per cent were of Borderline Intelligence, 21.9 per cent of Dull Normal Intelligence, 17 per cent Average Normal, and 1.9 per cent Superior. The Average Intelligence Quotient of patients under 16 years of age was 78, and of adults it was 73. An Average Intelligence Quotient of 77 was found for the idiopathic group and of 69 for the organic group. No significant sex difference was discovered.

Results of Our Study of Epileptic Group

The group studied by us was one that allowed the use of the Stanford Revision of the Binet tests as an appropriate measuring scale in most cases. A small group not scoring up to the three-year mental level was given supplemental items from the Kuhlman-Binet scale at the two-year level. Below this level, approximations based upon B scoring items of the Gesell Normative Schedules of Development were used. Only 11 of 103 cases examined fell in the group requiring supplementing of the Stanford-Binet below the three-year level. In addition to the Binet test, performance tests, educational tests, and other appropriate tests were used to fit the clinical needs of the particular case. This report is, however, essentially a report of the Intelligence Quotients of a literate group falling within the age range of the Stanford-Binet test. Retests were made of a limited group. This part of the paper deals with the results of the first test of the entire group. The results of retests will be reported later.

Table 15, which follows, gives the distribution of IQ's of the group. Table 16 shows the distribution and relation of chronological ages to mental ages of the same group. Table 17 introduces Dawson and Conn's epileptic group and presents a comparison of the intelligence of the epileptic group with that of two control groups in the same community; namely, the Los Angeles City school group (19, p. 6) and a group of children from the Children's Hospital suffering from allergy, mainly asthma (35). Table 18 presents in more detail a comparison with the two control groups.

TABLE 15

INTELLIGENCE QUOTIENTS OF CASES STUDIED

IQ	Number of cases
135-144	3
125-134	2
115-124	7
105-114	16
95-104	18
85- 94	21
75- 84	14
65- 74	3
55- 64	3
45- 54	5
35- 44	6
25- 34	2
15- 24	2
5 -14	1
Total	103
Mean	88
Median	92.4
Q_3	106.4
Q_1	60.1
Range	11 to 141

The Los Angeles City School data (19, pp. 6, 7) were chosen as representative of a more appropriate control group for this study than the Terman data (36, pp. 65-67). The allergy group was used as representative of a common basis of selection socially and economically as the epileptic group and as having the added advantage of representing the sick children group. The facts brought out in Tables 15, 16, 17, and 18 are as follows:

1. Our epileptic group, described in terms of central tendency (Mean or Median), is lower in intelligence than the City School group or the Allergy group. The Median Intelligence Quotient of our Epileptic group is 92.4; of the Los Angeles School group, 105; and of the Allergy group, 103 (Table 18).

TABLE 16
CHRONOLOGICAL AND MENTAL AGES OF CASES STUDIED

Age	Chron. age No. of cases	Mental age No. of cases
15	0	1
14	1	2
13	1	2
12	8	1
11	11	4
10	10	12
9	12	13
8	6	9
7	13	6
6	10	10
5	5	3
4	16	17
3	5	12
2*	5	8
1	0	3

Summary

Total:	103 cases					
Chron. age:	Mean	7 yrs. 6½ mos.	Mental age:	Mean	6 yrs. 6 mos.
	Median	7 yrs. 4 mos.		Median	6 yrs. 4 mos.
	Q_3	10 yrs. 0 mos.		Q_3	9 yrs. 2 mos.
	Q_1	4 yrs. 6 mos.		Q_1	3 yrs. 8 mos.
	Range	1 yr. 10 mos. to 14 yrs. 9 mos.		Range	9 mos. to 15 yrs.

*Age 2 years—interval 1 year 6 months to 2 years 5 months, etc.

TABLE 17

COMPARISON OF INTELLIGENCE IN EPILEPTIC GROUPS AND CONTROL GROUPS

Classification	IQ	L. A. City school percentage	Children's hospital allergy percentage	Children's hospital epileptic percentage	Dawson & Conn epileptic percentage
Superior	110	36.5	31.2	19.3	6.1
Average	90-109	47.9	62.1	36.8	20.3
Dull normal	80- 89	10.6	6.6	19.4	24.4
Borderline	70- 79	3.8	0.0	5.8	26.5
Feeble-minded	Below 70	1.3	0.0	18.4	22.4
Number of cases		63,147	45	103	49

TABLE 18

COMPARISON OF THE INTELLIGENCE QUOTIENTS OF THE EPILEPTIC GROUP, THE ALLERGY CONTROL GROUP, AND THE LOS ANGELES CITY SCHOOL CONTROL GROUP*

Intelligence Quotients	L. A. city school Percentage	Children's hospital allery Percentage	Children's hospital epileptics Percentage
Above 140	1.2	0	.9
130-139	3.1	4.4	1.9
120-129	10.2	8.8	6.8
110-119	21.8	18.0	9.7
100-109	27.3	26.6	19.4
90-99	20.6	35.5	17.4
80-89	10.6	6.6	19.4
70-79	3.8	0	5.8
Below 70	1.3	0	18.4
Number of Cases	63,147	45	103
Median	105	Mdn 103	Mdn 92.4
Other data not published		Q 8.9	Q 23.3
Can be approximated from above.		Range 81-139	Range 11-141

*The Los Angeles School data is selected from reports of Intelligence levels of Normal Six-Year Elementary School Children. (19, pp. 6, 7, Tables 2 and 3).

2. There is a greater variability in the intelligence of the epileptic group than in the School group or the Allergy group. The Quartile Deviation of our epileptic group is 23.3, of the Allergy group 8.9, and of the Los Angeles school group, not given but by approximation, is seen to be about one-half that of the epileptic group (Table 18).

3. The total range of intelligence of our group of epileptics approximates that of the non-selected school group. The epileptic group includes all *ranges of intelligence from very superior to very* inferior (Tables 15, 16, 17, 18).

4. There is a greater proportion of cases at the feeble-minded end of the scale in the epileptic groups (Children's Hospital and Dawson & Conn) than in either of the control groups (18.4 per cent in Children's Hospital group, 22.4 per cent in Dawson and Conn's group, and 1.3 per cent in Los Angeles City School group).

5. *In addition* to the large number of feeble-minded epileptics, the group includes a significantly large number of cases above the average or superior in intelligence. This group should be given consideration in understanding and treating epileptics.

6. The analysis of tests failed by the epileptic group as a whole does not show qualitative differences in responses from those given by a non-selected group. There are of course included the responses of feeble-minded epileptics characteristic of the feeble-minded and, of superior epileptics, responses characteristic of superiority in whatever group tested. There were observed no typically epileptic responses. Fox (12, p. 247) noted also in his group the typically feeble-minded response of his feeble-minded epileptics. His group included few of superior intelligence which accounts for his failing to note the superior responses.

7. One difference not evidenced statistically is recorded clinically, namely, the need of breaking the test period in order to avoid over-fatigue or other physical upsets. This was necessary in some cases, but not more noticeable in this group than in other sick children groups. Only one case of convulsion occurred in the psychological test period during the course of this work. As a precaution, we did not test children until the depressing after-effects of attacks had disappeared.

Our data do not show the degree of deterioration of the group reported by Fox (12). The Median for our group was 92.4 and for his group, 71 for Males and 65 for Females. Our group is also less deteriorated than the group studied by Dawson and Conn (5). The Mean for their group was 80.65 and for our group 88.0. The range for Dawson and Conn's cases was 49-117, our range, 11-141. Their Borderline and Feebleminded group comprise 48.9 per cent of the total group, whereas, our Borderline and Feebleminded group comprise only 24.2 per cent of the total group. This difference is possibly due to some different basis

of selection of cases for treatment in the hospital or a difference in the community furnishing the cases. Our results were in agreement with Dawson and Conn's conclusion that "the Mean IQ of epileptic children was appreciably and significantly below that of the other hospital children examined." The data of Dawson and Conn also show the same piling up of a large percentage of cases at the Feeble-minded end of the scale.

Variations in Intelligence Quotients of Epileptic Children Retested

Studies of the variability of the Intelligence Quotient of unselected children from the general population have shown that the IQ in these cases has a reliability high enough to make the results of one test have reasonable predictive value. Such studies are also of value for comparative purposes in relation to other more selected groups. Particularly is it interesting to compare such a group as the epileptic in variability with an unselected sample. Many studies have been made of this variability, some of which are to be discussed here. Our data included retests made at varying intervals with the epileptic group. The results of our treatment of these data will follow the discussion of previous studies.

Fox (12), Patterson and Fonner (28), Dawson and Conn (5), and Fetterman and Barnes (10) have investigated the variations in the intelligence of epileptic children by means of retesting after various time inter-

vals. Fox (12, p. 238), using Burt's revision of the Binet test, retested his group of 130 institutional epileptic children after an interval of approximately one year. Disregarding change in IQ of two points or less, he found that 37 per cent showed a decrease in IQ, 22 per cent showed an increase, and 41 per cent remained stationary. Eleven subjects showed a loss of over 10 points, whereas only one subject showed a gain of over 10 points. He interprets his results of retesting as showing a wide variation in IQ, the general tendency towards deterioration, and the very marked deterioration in over 8 per cent of the total.

Patterson and Fonner (28, pp. 542-548) carried out an extensive investigation of the variations of the IQ's of a group of 128 institutional epileptic children and adolescents (63 boys, 65 girls), in which retests were given at intervals of a year or more. Ninety-eight cases (51 boys, 47 girls) were tested twice by means of the Stanford-Binet Test; 30 cases (12 boys, 18 girls) were tested three times. Deviation of two points or less in IQ from the original test was disregarded and interpreted as showing no change. Deviation greater than two points was interpreted as showing an increase or decrease in IQ. Their results may be summarized: Of the 51 boys tested twice, after intervals ranging from one year four months to three years seven months, 13 cases showed no change on retest, 20 showed a decrease, and 18 showed an increase. The range of IQ on the original test was from 38 to 114, and on the second test from 45 to 113. The maximal loss in retesting was 17 points and the maximal gain was 18

points. Of the 47 girls tested twice after intervals ranging from one year zero months to three years 10 months, 18 cases showed no change in IQ, 13 showed a decrease, and 16 showed an increase. The range of IQ's on the original test was from 47 to 109 and on the retest from 42 to 109. The maximal gain was 26 points and the maximal loss was 9 points.

Of the 12 boys tested three times, eight were morons and four were dull normals. On the second test, two cases showed no change in IQ, five, a decrease, and five, an increase. On the third test, three showed no change from the original test, four showed a decrease, and five an increase. Of the 18 girls tested three times, 14 were mentally deficient (13 morons, one imbecile), two were dull normals, and two normals. On the second test, 10 showed no change in IQ, six showed a decrease, and two, an increase. On the third test, six showed no change, eight showed a decrease, and four an increase. Patterson and Fonner conclude that the IQ of epileptic children and epileptic adolescents shows a considerable variation from time to time, which may involve either an increase or decrease, and that such variations may occur in either direction. And, furthermore, the variations occur at any mental level, that is to say, patients of high and low IQ's show about the same rate and degree of fluctuations.

Dawson and Conn (5, pp. 145-147) retested (by means of the Binet Test: Burt's revision) 21 cases of his total group of 49 epileptic children. The time interval between tests ranged from eight months to four years eight months. The mean IQ of the first

test was 82.09±2.51, and of the second test 66.52±3.06. The difference, 15.57±3.96, shows a significant loss for the group as a whole. They conclude that "in ten cases there is little difference between the results of the two tests, but the general tendency is towards deterioration which in some is very serious. This supports the suggestion already made that the epileptic condition tends to produce mental deterioration" (p. 145).

Fetterman and Barnes (10, p. 799) tested twice or more, using the Stanford-Binet Test, 46 epileptic patients (ages not given) of his hospital dispensary group. The average time interval between tests was approximately 21 months. On first retest, 19 of these cases showed "a slight increase in intelligence quotient," 23 showed a "moderate loss," and in the remaining four cases the IQ was unchanged. They state, "The average of these changes is not larger than the difference which one may obtain between tests and retests on normal persons." Twenty-five of the 46 cases were retested three or more times. They say that "the most interesting feature of their serial studies is the fluctuation of the intelligence quotient from one test to another. What may look like a significant change one year may be entirely offset the next year by a change of equal magnitude in the opposite direction." For the majority of these 25 patients there was no decided trend toward mental deterioration in the Binet rating.

Our Results with Retests of Epileptics. Forty-four of our group of 103 cases were retested at intervals varying from one month to four years 11 months. The

average interval between tests was 14 months. Except in four cases, with an interval between the tests of two years five months to four years 11 months, the basis of selection for retests was availability in a routine procedure which aimed to secure retests on as many cases as possible. The four exceptions mentioned above were referred by the physician for retest because of suspicion of change of intelligence. Tables 19 and 20 show the differences in Intelligence Quotients on first and second tests.

TABLE 19

COMPARISON OF INTELLIGENCE QUOTIENTS OF EPILEPTICS ON FIRST AND SECOND TESTS

Intelligence Quotients	First test	Second test
135-144	1	0
125-134	1	3
115-124	2	3
105-114	7	9
95-104	11	5
85- 94	5	6
75- 84	6	5
65- 74	1	4
55- 64	2	0
45- 54	2	4
35- 44	4	3
25- 34	1	2
15- 24	1	0

No. of cases 44

First test:		Second test:	
Mean	86.7	Mean	87.2
Mdn.	95.0	Mdn.	91.6
Q_3	105	Q_3	109.5
Q_1	75	Q_1	70
Range	24-141	Range	26-130

TABLE 20

CHANGES IN INTELLIGENCE QUOTIENTS ON FIRST AND SECOND TESTS

Intervals between tests	No. of cases	Mean change in points	Direction of change: Pos.	Direction of change: 0	Direction of change: Neg.	Mean change in direction in points	Range of change in points
1 to 5 mos.	4	3.5	3	0	1	+ 2.0	— 3 to + 5
6 to 11 mos.	11	6.3	5	0	6	+ 1.4	—10 to +19
12 to 17 mos.	19	7.6	8	2	9	— 0.1	—27 to +24
18 to 23 mos.	6	4.8	4	0	2	+ 3.3	— 1 to +11
24 to 35 mos.	2*	12.0	0	0	2	—12.0	— 3 to —21
36 mos. and up	2*	27.0	0	0	2	—27.0	—26 to —28
Number of cases			44	Mean change in points			7.6
Changes in positive direction			20	No change			2
Changes in negative direction			22	Mean change in direction			— .83

*These cases were referred for retests because of suspected deterioration and consequently cannot be regarded as typical. If they were omitted in our calculations, the change in the group would be more typical of changes in a non-selected group.

Table 19 shows a slight increase in the Mean IQ of the second test. This difference, +.5 points, is not statistically significant. The difference in the case of the Median is a decrease on the second test of 3.4 points. In fact, the Median and 25th percentile show a general lowering of the group. Table 20 shows a large negative change in the cases tested after an interval of 23 months. However, this group includes only four cases who, as already stated, were selected for retests by the physician on the basis of suspected deterioration and from them no general conclusions can be drawn. The cases do, however, tell the story for these particular cases even though they are not an adequate sample of the group. Table 20 shows an increase in IQ in 20 cases, a decrease in 22 cases and no change on retest in two cases. The Mean change in IQ points on the second test is 7.6, the Mean change in direction —.83 IQ points. The range includes changes for +24 to —28 points. The middle 50 per cent change was from 4.5 points increase to 4 points decrease.

The Pearson Product Moment correlation between IQ's on first and second tests was +.897 (P.E. of *r*, ±.019). Terman's data on retests (37, pp. 137-164) show a higher correlation on retests of a non-selected group. His correlation was +.933 (based upon 435 cases retested two or more times, through ages three to 15 years with intervals of one day to seven years). A comparison of our epileptic group with Terman's non-selected group of school children shows greater variability on retests of the epileptic group.

We reorganized the data to represent the IQ not in points but in terms of mental classification as defined in Terman's handbook (36) namely: Superior, Average Normal, Dull Normal, Borderline, Feeble-minded. This treatment showed the following results with the 44 cases retested: no change in diagnosis on retest in 29 cases, change to higher classification in seven cases, change to lower classification in eight cases. The changes were as follows: Average to Superior, three cases; Dull to Superior, one case; Dull to Average, one case; Borderline to Dull, two cases; Superior to Average, three cases; Average to Dull, one case; Average to Borderline, two cases; Dull to Borderline, one case; and *Average to Feeble-minded, one case.* The last case was the only one which brought about a Feeble-minded diagnosis on the second test and not on the first test. A point of greater interest is that not one of the cases diagnosed as Feeble-minded on the first test improved enough to be given a higher classification on a second test.

In seven cases having three tests the Mean IQ on successive tests was as follows: First test, Mean IQ 86.2; Second test, Mean IQ 81.1; Third test, Mean IQ 82.5.

In one case tested five times at approximately four years, eight years, nine years, 11 years, and 12 years, the IQ's were as follows: 95, 67, 69, 87, 77. This case, which was one of Jacksonian epilepsy, shows deterioration after the first test and great variability.

A special statistical treatment was given to a selected group tested who were diagnosed as suffering

from idiopathic epilepsy. This group retested included 22 cases, examined at intervals from one month to two years, five months. The Mean IQ of this group on the first test was 87.2, on the second test 85.8. This shows a difference of Means of -1.4 ± 3.0. Although this is not statistically significant it is in agreement with other results reported which indicate a slight trend downward of the groups studied.

Our retest data dealing with cases diagnosed as organic in type are too small to be given statistical treatment. It includes five cases with an Average IQ on first test of 95 and on second test, of 96 (10 months to four years 11 months later). The individual variability in these cases was great, including a change from Normal to Feeble-minded. The changes in IQ's were as follows: 114 to 111, 82 to 92, 112 to 131, 74 to 82, and 95 to 67.

For clinical purposes we are adding a brief analysis of the cases having the greatest change in intelligence level (including cases changing 10 points or more in either direction). These include seven improving and six deteriorating cases.

Improving Cases.

Case No. 68: Improved 34 points (Dull Normal to Superior). Male. Major epilepsy and equivalent. Onset at three months. Birth history adverse (blue baby). Idiopathic. Neurological findings negative.

Case No. 1: Improved 30 points (Average to Superior). Female. Major epilepsy. Onset at 10 years, four months. Birth history adverse (premature, immature). Idiopathic. No neurological examination.

Case No. 49: Improved 19 points (Superior to More Superior). Female. Minor epilepsy. Onset at five years. Birth history shows birth injury. Organic Jacksonian. Neurological findings include: spastic paraplegia, choreoathetosis, adductor spasms, deep reflexes increased, bilateral Babinski.

Case No. 3: Improved 14 points (Average to Superior). Female. Major first, later minor. Onset at one year two months. Birth history normal. Idiopathic. No neurological examination.

Case No. 89: Improved 12 points (Average to Superior). Male. Major epilepsy. Onset at six years six months. Adverse birth history (premature). Organic. Neurological findings: deep reflexes increased.

Case No. 85: Improved 11 points (Average to Improved Average). Male. Major epilepsy. Onset 10 months. Adverse birth history (precipitate birth). Organic? Neurological examinations incomplete; indicates pathology.

Case No. 11: Improved 10 points (Dull to Average). Male. Major and Minor. Onset at four years five months. Adverse birth history (long labor). Organic. Neuro-pathology recorded. (Jacksonian Convulsions.)

Deteriorating Cases.

Case No. 84: Loss of 28 points (Average to Feeble-minded, 95 to 67 IQ). Female. Minor epilepsy. Onset at two years. Birth history normal. Organic. Findings of neurological examination: Spastic right hemiplegia, choreoathetosis, muscular atrophy of right side, deep reflexes increased, right Babinski.

Case No. 62: Loss 27 points (Superior to less Superior). Male. Major epilepsy. Onset at 11 months. Adverse birth history (three days labor). Idiopathic. Neurological findings negative.

Case No. 81: Loss 26 points (Average to Borderline). Female. Major epilepsy. Onset at six months. Birth history normal. No neurological examination. Classified Organic (inconclusively).

Case No. 14: Loss 21 points (Average to Borderline). Female. Major and Minor epilepsy. Onset at four months. Birth history normal. Idiopathic. Neurological findings state negative but report "hypertonia."

Case No. 28: Loss 18 points (Superior to Average). Male. Major epilepsy. Onset at one year three months. Birth history normal. Idiopathic. Neurological examination negative.

Case No. 10: Loss 10 points (Low Dull to Borderline). Male. Major epilepsy. Onset at six years six months. Birth history normal. Idiopathic. Neurological examination negative.

It is difficult to see any significant difference in the causative factors accompanying epilepsy in the cases showing most improvement and those having greatest deterioration. The idiopathic classification occurs four times out of five cases showing a large amount of deterioration, and in only two out of six of the improving cases. The neuropathology described in the case deteriorating the most (a loss of 28 points) is very similar to the description in a case that improved a great deal (gained 19 points). However, the greatest changes occurred in cases classified above the level of 70 IQ (not feeble-minded). Throughout the greatest irregularity and variability in performance at different times is shown by individuals of average or superior intelligence. Seven of 13 cases varying most had a classification as superior on either the first or second test.

The following is a brief summary of significant facts brought out in our data with reference to variability of intelligence quotients of epileptic children.

1. The epileptic group shows greater variation on retests than is reported in studies of unselected school children.

2. There is shown mental deterioration of the group as a whole on retests.

3. Our group does not show as great a deterioration on retests as is shown by Dawson and Conn, Fox, and others earlier in this report.

4. Our results show no large changes that would change the diagnostic classification in 65.6 per cent of the group. Changes of 10 points or more occur in 34.4 per cent of the cases retested. These changes are in both directions (improving and deteriorating).

5. Large changes in a positive direction or a negative direction occurred most frequently in cases of Average or Superior Intelligence.

6. No change in a positive direction occurred in an amount great enough to result in a child classified as Feeble-minded on a first examination, later to be classified as not Feeble-minded.

7. In only one case did a child classified as normal on a first examination deteriorate so rapidly as to be classified as Feeble-minded on a later test.

SEX AND INTELLIGENCE IN EPILEPSY

Our data show that, in the total group of epileptics examined, the males were present in larger numbers than the females and that the actual proportion of

males in the epileptic group was greater than in the general hospital population of new cases.

Slight differences are shown in the intelligence levels of male and female epileptics. The group included 63 males and 40 females. On the first test the males had a Mean Intelligence Quotient of 86.3 and the females 90.5. The Mean Intelligence Quotient for the entire group was 87.9. This shows the female epileptics less deteriorated as a group than the males.

Twenty-nine males were retested and 15 females. The Mean Intelligence Quotient of the 15 females retested was 88.6 on the first test and 87.6 on the second test. The Mean Intelligence Quotient of 29 males retested was 85.6 on the first test and 86.3 on the second test. The results show a slight improvement in the male group not statistically significant and a slight loss in the female group also not significant. Here again, although not shown in the Mean, the total group shows great individual variability but no characteristic differences between males and females in variability.

Severity of Attacks and Intelligence

The literature is limited and the data inconclusive with reference to the relationship of severity of attacks to level of intelligence.

Dawson and Conn (5, p. 151), Doolittle (6, p. 96), Fetterman and Barnes (10, p. 801), and Worster-Drought (38, p. 60) state that the severity of attacks and subsequent mental condition are not significantly related. Grinker (15, p. 833), on the contrary, states

that mental deterioration bears some quantitative relation to severity of attacks.

Our data afforded opportunity to study this problem. We found that the Mean Intelligence Quotient in 55 cases having major convulsions was 87.2 and the Mean Intelligence Quotient of 16 cases having only a minor epilepsy was 97.0. The difference appears to have some significance; however, the small number of cases in the group of minor epilepsy leads us to present this difference as suggestive rather than conclusive.

CAUSATION AND INTELLIGENCE

We mentioned earlier some facts with reference to intelligence level in cases of idiopathic and organic epilepsy. We found in the case of 53 epileptics with idiopathic epilepsy a Mean Intelligence Quotient of 88 and in the case of 18 epileptics with an organic diagnosis a Mean Intelligence Quotient of 79.8. Here again, the variability of the groups, the limited number of cases, and the incompleteness of the neurological data lead us to present this as inconclusive evidence of lower level of intelligence in the cases of organic epilepsy than in idiopathic epilepsy.

INTELLIGENCE AND AGE OF ONSET

Gowers (14, p. 122 f.) found that the age of onset of epilepsy has a definite influence on the mental condition of the patient. Thus, of 64 cases with "considerable mental defect" 55 per cent of these had their onset under 10 years of age, whereas of his whole

series of 1450 cases, only 29 per cent had onset before 10 years. Grinker (15, p. 833), Paskind (26, p. 385), and Worster-Drought (38, p. 60) all assert a relationship of mental deterioration to early onset.

In our study the age of onset was determined by conferences with parents and study of the social and medical histories. It was able to be determined in 102 cases. A tabulation of data showing age of onset in relation to Intelligence was made. No significant relationship was shown when the change of Intelligence Quotients was studied with reference to small intervals (six months or one year). However, when grouped in larger intervals, a difference was indicated.

In the cases examined (72) having onset of convulsion from birth to five years 11 months, the Mean Intelligence Quotient was 82.87. In the cases having onset from six years to 12 years (30), the Mean Intelligence Quotient was 90.43. The difference (7.56 points) is statistically significant. This shows that the cases having an early onset are more deteriorated as a group than those having an onset later in childhood. The difference of Means does not indicate that all cases are deteriorated directly in relation to age of onset but that this is one factor in deterioration of intelligence.

Table 21 shows details of facts of age of onset and intelligence of cases tested. This table also shows the early onset in a large number of cases.

Frequency of Attacks and Intelligence

Our data do not allow for the study of relation of

TABLE 21
AGE OF ONSET AND INTELLIGENCE

Age of onset in months	Number of cases	Average IQ
144-155	1	105.0
132-143	3	95.6
120-131	3	97.0
108-119	4	72.7
96-107	5	94.8
84- 95	5	88.4
72- 83	10	106.2
60- 71	5	90.4
48- 59	5	90.2
36- 47	13	75.9
24- 35	10	96.2
12- 23	17	90.7
0- 11	21	76.0
Total	102	

Mean Intelligence Quotient: 90.43 (6 yrs. to 12 yrs.) 72 cases.
Mean Intelligence Quotient: 82.87 (0 yrs. to 5 yrs. 11 mos.) 30 cases.

frequency of attacks to intelligence level or personality changes due to lack of complete data on frequency in cases studied.

The problem as to the relationship of frequency of attacks to intelligence has not been settled. Patterson and Fonner (28, p. 548) state that the variations in the intelligence of epileptics seem in the main to be independent of the frequency of attacks. Dawson and Conn (5, p. 151) reported no significant correlation between frequency of attacks and subsequent mental progress. Stevenson (34, p. 218) reports that the frequency of attacks and mental deterioration are related. Of his series of organic cases (with severe head wounds), 17 of 17 cases (i.e., 100 per cent) with frequent attacks were deteriorated, whereas 23 of 37 cases (i.e., 62 per cent) with infrequent attacks were deteriorated. Of

his 108 cases of idiopathic epilepsy, 94 per cent of those with frequent attacks were deteriorated, as compared with only 47 per cent of those with infrequent attacks. Doolittle (6, p. 96) found no definite relationship between frequency and rapidity or certainty of mental deterioration. Paskind (26, p. 385), on the other hand, found that non-deteriorated patients had on the whole less frequent attacks than the deteriorated ones. Grinker (15, p. 833) says that frequency of attacks bears a relationship to mental deterioration. Worster-Drought (38, p. 60) asserts that frequency and mental deterioration are on the whole unrelated.

Duration and Mental Deterioration

The deteriorating effect of continued convulsions on the intelligence of the individual has been noted. Worster-Drought (38, p. 58) says that "the more pronounced mental characteristics of epilepsy are usually seen when the disease is of rather long standing." Doolittle (6, p. 96) states, however, that the duration of convulsions has no definite bearing upon the rapidity or certainty of mental deterioration. Fetterman and Barnes (10, p. 801) likewise found no direct relationship between intelligence quotient and duration of epilepsy.

We selected for study 73 cases that were continuing to have seizures at the time covered by the examinations (within one month of examination). A study of these cases showed a direct relation of intelligence level and continuation of the epilepsy. Table 22 summarizes the data on duration and intellectual deterioration.

TABLE 22
RELATION OF INTELLIGENCE TO DURATION OF EPILEPSY

Examination after epilepsy present for following intervals	Number	Mean IQ	Range
Less than 6 mos.	13	100	82-125
6 mos to 11 mos.	9	105	59-125
1 yr. to 1 yr. 11 mos.	11	94	24-141
2 yrs. to 2 yrs. 11 mos.	13	91	62-135
3 yrs. to 5 yrs. 11 mos.	14	86	32-112
6 yrs. to 11 yrs. 9 mos.*	13	68	11-107

*Eleven years, 9 months represented limit of duration of epilepsy in cases studied.

Even though the results are conclusive with reference to relation of duration to deteriorating Intelligence Quotient in the cases studied by us, it will have to be remembered that the cases with a long duration in the children's group are also cases of early onset. (Early onset was mentioned earlier as a factor related to deterioration.)

Also the range in each group is important. The average is much affected by badly deteriorating cases, and each group includes cases that do not seem to deteriorate a great deal even with long duration of the epilepsy. Here the variability of a group is of as great significance for individual prognosis as is the Mean for the group.

FREEDOM FROM ATTACKS AND INTELLIGENCE LEVEL

The literature included little reference to cases having onset of convulsions and later cessation. Our data included on tests and retests the following types of cases: children continuing to have convulsions in the

interval covered by the psychological examination, and children who had for varying intervals just preceding the examination been free from convulsions or other epileptic attacks.

Of the total 103 cases studied, 75 were having convulsions at the time of the psychological examination and 28 had been free of attacks for a period of from one month to five years two months.

The Mean Intelligence Quotient for the group of 103 was 88, the Mean Intelligence Quotient for the 75 continuing to have convulsions was 91.2.

Twenty-three of the 75 cases having convulsions at the time of first test continued convulsive through a second test period. The Mean Intelligence Quotient of this group, which is a group not improving, was 80 on first test and 81 on second test.

Seventeen cases having convulsions at the time of first test had an interval of freedom from convulsions before retest (four months to one year two months.) This group of improving cases had a mean Intelligence Quotient of 99 on the first test and 98 on the second test.

Thirty-eight cases tested after varying intervals of freedom from attack had a Mean Intelligence Quotient of 83.6. This group included: 10 cases free three to five months, 10 cases free six months to 11 months, 12 cases free one year to one year 11 months, two cases free two years to two years 11 months, and four cases free three years to seven years.

The above data are somewhat contradictory. However, the following conclusions seem justified:

1. Cases having convulsions at the time of first test are slightly higher in intelligence than the total group which includes some improved cases.

2. Cases having convulsions and found later not to improve in this respect are lower in intelligence than the convulsive cases that later improved. The group that improved, although having a higher general level of intelligence, did not show improvement or deterioration in intelligence on a second test when convulsions had ceased.

The cases all seemed to maintain their mental status as groups regardless of duration of seizures or interval of freedom from convulsions. The improvable cases were originally less deteriorated than those that did not improve.

3. In general the group having a history of convulsions that later ceased were lower in intelligence level than the total epileptic group and particularly than the group having convulsions at the time of first or second examinations. This again may be related to early age of onset and not freedom from attacks.

IV

DISCUSSION OF THE SO-CALLED "EPILEPTOID PERSONALITY"

The existence of a specific epileptoid personality has been frequently asserted. Notkin (25, p. 799), for example, says, "There has been a tendency in recent years to look on the personality traits of epileptic persons as a specific manifestation. These traits are egocentricity, supersensitiveness, irritability, emotional poverty and stiffness of mentation." Doolittle (6, p. 89) says that descriptions of the epileptic personality "vary somewhat as to its characteristics from that of simple irritability to the most severe forms of personality derangement, including supersensitiveness, hyperirritability, 'religiosity,' narcissism, uncontrollable outbursts of rage, furors and sexual assaults." Grinker (15, p. 833) writes that epileptics "are said to be egocentric, selfish, seclusive, and show tremendous outbursts of emotion. Their personality is said to be rigid. When they cannot have their way, they show a great amount of rage; children, especially become frenzied when refused their wishes." Worster-Drought (38, p. 59) says that the epileptic patient "is frequently egocentric, his thoughts and interests centering around his own personality." The asserted characteristics of the epileptoid personality are principally described as egocentricity, irritability, and outbursts of emotion.

Fay (9, pp. 557 f.) has found that, in the case of mentally deficient epileptics suffering from major convulsions, *the release of the patient from seizures by*

means of dehydration frequently gives rise to a behavior problem, "which is even more undesirable and difficult to manage than the condition formerly presented by frequent convulsive seizures." Eyrich (8, pp. 640-645) has described three typical syndromes of the personality of epileptic children, which syndromes are not specifically nor exclusively epileptic in meaning. The central point of the first syndrome is a retardation of all mental processes ["eine Verlangsamung der gesamten psychischen Vorgänge" (p. 641)]. This retardation shows itself in aperceptive capacity, thought, and activity. In extreme cases there is an enormous retardation of reaction time, impoverishment of acts, and a limitation or loss of spontaneous activity. This personality syndrome we may call the mentally retarded hypokinetic type. The second syndrome described by Eyrich is the syndrome of explosive excitability ["Syndrom der explosiven Reizbarkeit" (p. 642)], which is characterized by an egocentric hypersensitivity toward trespass upon the patient's interests and self value with tendency to emotional outbreaks, "motor abreactions," and ill-humor. This syndrome is found in children, adolescents, and adults. The third syndrome ["Symptomenbild einer hyperkinetischen Erkrankung im Kindesalter" (p. 643)] is characterized by an elementary, impulsive motor restlessness, without a goal. There is an increased distractibility, and with respect to emotion, nervous tension and a "genial vacuity." This last syndrome, which appears to be limited to childhood, has an unmistakable relation to the psychological change following epi-

demic encephalitis. Eyrich further states that the first two syndromes are more common than the third, and that all appear more frequently in gradations and combinations than in pure form, especially between the first two syndromes.

Kramer and Pollnow (16) have described a syndrome (not clearly differentiated from other nervous disorders), occurring in 45 children of whom 19 were convulsive. The principal feature of this syndrome is a constant restlessness of a purposeless nature, with great distractibility. This syndrome resembles the third one described by Eyrich with respect to hyperkinesis, but with reference to the emotional status, whereas Eyrich found nervous tension and "a genial vacuity," Kramer and Pollnow found frequently aggressive and negativistic traits and outbursts of rage. All but three cases showed disturbances of speech. The usual onset of this syndrome was between three to four years of age, although an earlier onset was not uncommon.

Notkin (25, p. 800) studied the personality of non-psychotic epileptics in a general hospital (New York Post-Graduate Neurological Clinic) and found "in the great majority of patients little or nothing of the so-called epileptic make-up." He further studied a selected group of 150 epileptic persons (75 men, 75 women) of the Manhattan State Hospital, concerning whom more or less reliable life histories were obtained, and again "was impressed with the frequent occurrence of normal types of personality" (p. 800). He reports that the age at onset has apparently a direct bearing

on the type of personality; when the onset appeared in infancy or early childhood, the personality was decidedly "epileptoid." Pure "epileptoid" traits, as heretofore described, were found in cases in which the onset of convulsions occurred before the age of 12 years.

Dawson and Conn (5, p. 150) interpret the epileptoid personality when present in terms of retardation of mental functions. According to them, "a general enfeeblement in intelligence will, in itself, to some extent, explain the inattentiveness, the distractibility, the slowness of thought and reaction, the ego-centricity and the primitive character of the emotional reactions of epileptics, for it implies a disintegration of the higher, more complex mental functions and a consequent slackening of control over the lower, more purely instinctive reactions."[11] The personality of the epileptic, according to Doolittle (6, p. 96), is greatly influenced by the isolated and protective environment of institutional life. Paskind (26, p. 373) states that 93.5 per cent of a series of 304 epileptics from private practice were in "excellent mental health," and the remaining 6.5 per cent were "deteriorated."

[11]This is in acordance with Hughlings Jackson's doctrine of deficiency and release phenomena as revealed by lesions of the central nervous system.

V

OUR OBSERVATIONS ON THE RELATION OF INTELLIGENCE TO PERSONALITY AND CONDUCT, AND THE NATURE OF THE PERSONALITY OF EPILEPTICS

We offer these data and the conclusions drawn therefrom as a preliminary report of a tentative nature. Many difficulties stand in the way of collecting data with reference to emotional, volitional, and ethical behavior, having even a reasonable degree of objectivity and reliability. If such data of a quantitative and qualitative descriptive type are collected with reference to a particular group such as the epileptic, there are no available norms of behavior of control groups. An added difficulty is that of segregating the particular causative factor to be studied in its relation to behavior. This is because of the numerous and complex factors influencing behavior and personality. They include factors in the environment such as: home, school, and community conditions, influences, and attitudes; other facts of physical development and physical health; special factors of education and training; and the effects of the above on the variety of developing personalities.

In spite of the difficulty, we have treated the available data quantitatively as it allows, and summarized certain descriptive facts shown. We present this as suggestive of further lines of study of the epileptic. The data are not of test type except for intelligence

tests. They include observations made during the test period by the psychologist, reports from physicians and other hospital staff members dealing with and observing the case over a period of months or years, conferences with parents; school reports, reports of social workers on conditions in the home, and, in a few instances, Court records. We attempted to get a description of behavior as objective as possible and in addition, a general classification of the individual as a problem of personality or conduct or not such a problem. The facts brought out by the statistical treatment of these data having reference to our epileptic children are suggestive. The following are some of the findings of this study:

1. Of the total 103 cases, 44 cases were classified as having serious and continued personality difficulties or conduct disorders, or both. Four more were questionable problems. Forty-five were considered well adjusted at home, at school, and in the Hospital. *They did not present problems of personality or conduct.* Ten cases did not have sufficient data to justify treatment.

2. Further analyses of the 44 problem cases showed 31.8 per cent female and 68.1 per cent male (the expected ratio in the epileptic group was 38.8 female to 61.2 male). This shows a slightly and possibly significantly greater number of boys presenting serious problems.

3. The Mean Intelligence Quotient of the group presenting problems was less than the Mean of the total epileptic group and necessarily less than the epileptic group not presenting problems of behavior. This was due in large measure to the larger percentage of feeble-minded in the group considered problems. These results can be best presented in tabular form (Table 23).

TABLE 23

INTELLIGENCE OF PROBLEM AND NON-PROBLEM CHILDREN IN EPILEPTIC GROUP

Problem cases	44	Non-problem cases	45
Mean IQ	75	Mean IQ	88
Median IQ	83	Median IQ	92.4
Q_3	96	Q_3	106.4
Q_1	50	Q_1	100.1
Q	23	Q	23.1

The number of Feeble-minded problem cases was 14 of a total of 19 Feeble-minded cases. The number of those not Feeble-minded (above 70 IQ) presenting problems was 30 of a total 84 in the group. The Feeble-minded thus contributed 31 per cent of the problem cases and those not Feeble-minded 68 per cent. (The expected ratio was 18.4 per cent feeble-minded to 81.6 per cent, not feeble-minded). Stating the above differently, 35.7 per cent of the epileptics above 70 IQ were problems, 73.7 per cent of those below 70 were problems.

Age presented no explanation of problem cases. The Mean Age of the problem cases was the same (seven years six months) as that of the total group. The problem cases included the total range of ages.

Birth history was not explanatory of behavior problems in the majority of cases. A normal birth history was reported in 22 cases, birth injury in four, adverse birth factors in seven. It is worthy of note that four of five birth injury cases were problems.

Severity of convulsive illness seemed not to distinguish the problem cases. Of 44 cases in the problem group, the proportion of Major to Minor epilepsy was

83 per cent to 16 per cent. In the total group the proportion was 81 to 18 per cent.

The idiopathic as a group was represented in the same proportion in the problem group as in the group as a whole; in the problem group 64 per cent were idiopathic and 35 per cent, organic; in the group as a whole, 63 per cent were idiopathic and 36 per cent, organic.

The onset of the epilepsy was typically at an earlier age in the cases of children who were problems of personality or conduct than in the case of those that did not present such problems. In the problem group, 27 per cent had onset of the epilepsy at one year or before that age; 34 per cent, two years or sooner; 50 per cent, three years or sooner; and 75 per cent, six years or sooner. In the non-problem group only 15 per cent had onset at one year or sooner; 21 per cent, two years or sooner; 38 per cent, three years or sooner; and 74 per cent, six years or sooner.

The problem group did not show more rapid deterioration or greater variability of Intelligence Quotient than the group as a whole. Seventeen of 24 problem cases retested showed no change in mental classification (Superior, Normal, Dull Normal, Borderline, Feebleminded); four improved their classification (Normal to Superior, one; Borderline to Dull Normal, one; Dull Normal to Superior, one; Normal to Superior, one); three cases dropped to a lower classification (Normal to Borderline, one; Dull Normal to Borderline, one; Normal to Dull Normal, one). The Mean increase in points on retests for the problem

group was +2.4 (slightly but not significantly better than the total group).

The above facts show that the problem group is a fair sampling of the group as a whole except for the greater proportion of the Feeble-minded.

The most significant and interesting results have to do with the analysis of the nature of personality and conduct disorders in the group, and particularly those problem cases within the epileptic group that are not Feeble-minded. For convenience we will refer to the group (Superior, Normal, Dull Normal, Borderline) as the Normal problem cases and those below 70 IQ as Feeble-minded problem cases.

One fact of significance in understanding behavior and personality problems is the home situation (physical, economical, mental, social, educational, ethical). The group of problem epileptics (both feeble-minded and normal) had almost without exception poor home environment. Bad conditions were of the same type and as generally present as have been many times reported in groups of delinquent children. Two factors were worse in this group than ordinarily reported in delinquent groups: namely, universal economic insufficiency and presence in the home, in many cases, of other behavior problems of mental and health type (epilepsy, insanity, feeble-mindedness, etc.).[12]

Problem behavior reported might be classified under the following general headings: Hyperkinetic, Emotional and Other Personality Maladjustments, Conduct Disorders.

[12]Cases of two or more of the same family suffering from epilepsy were noted (father, mother, and other children in the family).

In the Feeble-minded problem group, the hyperkinetic behavior was reported in 13 of the 14 cases; emotional disturbances in 10 cases; and conduct disorder in only 5 cases. The last tendency, namely, not to classify Feeble-minded children as conduct problems, indicates a common tendency to ignore ethical implications in case of the feeble-minded and to consider the behavior as a phase of the feeble-minded problem.

In the Normal Problem group (not Feeble-minded) hyperkinetic behavior was reported in 18 of the 30 cases, emotional disturbances in 22 cases, and conduct disorders in 24 cases.

The hyperkinetic behavior was characteristic of the problem group as a whole. It included such descriptive statements as: hyperactive, constantly into things, fidgets, ill at ease, and talkative.

Emotional disturbances also were characteristic of Normal and Feeble-minded problem cases. The emotional behavior of the Normal, however, was more varied with reference to conditioning situations and complexity of responses.

Emotional maladjustments were reported in such descriptive statements as: rage, anger, temper tantrums; cruelty to children and animals;[13] quarrelsome fights, nags; destructive, throws dishes, cuts things with scissors; irritable, impatient, high-strung; mischievous, teases, pinches, bites, annoys; wanders away,[14] runs

[13]One child nearly choked a dog to death. When fingers were released he said, "Couldn't help it."

[14]One boy in Court because of wandering away constantly.

away; doesn't get along with other children, shy, hypersensitive; paranoid, delusions; attempted self-injury, attempted suicide;[10] negativistic, rebellious, moody, stubborn, sulks;[10] demands attention, egocentric, center of stage, self-centered.

Conduct disorders reported included: bad toilet habits, bad sex habits, lying, stealing, fighting, disobedience, non-cooperation, naughtiness, destructiveness, and others characteristic of delinquent and socially maladjusted children in general.

The upset in developing habits was reported and is possibly more general than our data show: nail-biting, thumb-sucking, masturbation, and speech disorders of various types.

---

[10]Three cases (10 per cent) of self-injury and attempted suicide in 30 Normal Problem epileptics (one cut his arm with glass; one attempted to throw herself in front of car; one was stopped from throwing himself from a high window of a building).

[10]One boy stayed in a high tree until 11 P.M. and had to be brought down, because he was angry with his mother about some slight matter.

VI

SUMMARY

In summarizing our observation of behavior disorders in epileptic children we might note the following:

1. The disorders of personality and conduct are not characteristic of all children with epilepsy. A large group present normal personality and behave in a normal manner.

2. The feeble-minded epileptic is more often considered a problem than is the normal epileptic child, possibly not because of a greater amount of epileptic type of behavior, but because his typically feeble-minded response makes him a problem.

3. Those individuals in the epileptic group classified as problems present common types of problem behavior namely, hyperkinesis (restlessness, hyperactivity, talkativeness) and various forms of emotionally upset behavior (irritability, temper tantrums, moodiness, egocentricity, hypersensitivity and shyness).

4. *Conduct disorders typical of children in good* health with bad environmental background and unsatisfactory training are found in the behavior reported in the case of problem epileptics; namely, lying, stealing, fighting, sex misbehavior, cruelty, destructiveness, etc.

5. We cannot answer finally the question of the existence of a typically epileptoid personality and behavior. However, if and when an epileptoid personality exists, it may have a variety of explanations. The most likely partial explanations seem to be these:

a. Epilepsy and the epileptoid personality may both be the expression of an underlying constitution (the "constitutional" explanation).

b. The disease process which involves epilepsy may alter the personality so as to make it epileptoid (the "disease-altering-personality" explanation). That is to say, the epileptic condition imposes itself on the personality and produces modifications therein. The resulting epileptoid personality is a derivative of the personality existing before epilepsy.[17]

c. The invalidism and isolation, physical and mental, of epilepsy, particularly the self-knowledge of one's condition, may produce the epileptic personality. Also other social, economic, mental, and educational facts may condition and direct development of abilities, habits, and attitudes (the "environmental explanation").[18]

d. Epilepsy may develop on the basis of a preexisting epileptoid personality (the "psychogenic" explanation).[19]

---

[17]Post-encephalitic and post-traumatic behavior problems have been frequently described. Drayton (7), as already stated, gives evidence that post-traumatic changes in behavior are organically determined.

[18]Bridge (2, p. 724) says that "in order to gain insight into the origin of the peculiarities of temperament of the epileptic patient one must know the family background and social setting in sufficient detail to be able to make distinctions between cause and effect." And further he says (p. 735), "In all probability the so-called epileptic personality is not an entity which bears any specific causal relationship to the disease, but represents, in large part, the response of such patients to the problems and situations which the very nature of their disease creates."

[19]The rôle of emotion in precipitating attacks in established epilepsy is well known. Fremont-Smith (13, p. 234 f.) for example

e. The epileptoid personality may in reality be the mentally retarded behavior existing in mentally retarded epileptics (the "mental retardation" explanation).

Our results give some support to each of the above explanations. The problem remains one of isolation and study of behavior with each factor controlled.

6. It is not possible to distinguish in the epileptic group of normal intelligence the cases that will become problems and present the hyperkinetic behavior and emotional upset that seem to characterize the epileptic problem cases. We found no significant relation to age, sex, causation, or severity. We found more cases in the problem group had an early age at onset of the epilepsy than did those in the non-problem group. This indicates some relation between personality and conduct problems and early age of onset or duration of the illness or both.

---

found that in 24 of 36 patients of private practice, whose ages ranged from 10 to 53 years, and who suffered major convulsions, emotion was a precipitating factor in producing attacks. In eight of these cases, major or minor seizures were precipitated by a discussion of the emotional difficulty involved. Both organic and idiopathic cases were involved.

References

1. Bentley, M., & Cowdry, E. V. The problem of mental disorder. New York: McGraw-Hill, 1934. Pp. x+338. (See esp. Section II, Chapter 2 by A. Myerson, 28-38.)
2. Bridge, E. M. Mental state of the epileptic patient. *Arch. Neur. & Psychiat.*, 1934, 32, 723-736.
3. Cobb, S. Causes of epilepsy. *Arch. Neur. & Psychiat.*, 1932, 27, 1245-1263.
4. Cookson, S. H. An analysis of one hundred cases of fits in children. *Arch. Dis. Childhood*, 1927, 2, 178-185.
5. Dawson, S., & Conn, J. C. M. The intelligence of epileptic children. *Arch. Dis. Childhood*, 1929, 4, 142-151.
6. Doolittle, G. J. The epileptic personality, its progressive changes among institutional cases. *Psychiat. Quar.*, 1932, 6, 89-96.
7. Drayton, W. Pneumocranium in the treatment of traumatic headache, dizziness and change of character. *Arch. Neur. & Psychiat.*, 1934, 32, 1302-1309.
8. Eyrich, M. Über Charakter und Charakteränderung bei Kindlichen und Jugendlichen Epilepikern. *Zsch. f. d. ges. Neur. u. Psychiat.*, 1932, 141, 640-645.
9. Fay, T. Convulsive seizures, their production and control. *Amer. J. Psychiat.*, 1931, 10, 551-566.
10. Fetterman, J., & Barnes, M. R. Serial studies of the intelligence of patients with epilepsy. *Arch. Neur. & Psychiat.*, 1934, 32, 797-801.
11. Foerster, O., & Penfield, W. The structural basis of traumatic epilepsy and results of radical operation. *Brain*, 1930, 53, 99-119.
12. Fox, J. T. The response of epileptic children to mental and educational tests. *Brit. J. Med. Psychol.*, 1924, 4, 235-248.
13. Fremont-Smith, F. Influence of emotion in precipitating convulsions. *Arch. Neur. & Psychiat.*, 1933, 30, 234-235.
14. Gowers, W. R. Epilepsy. London: Churchill, 1881. Pp. xiv+309.
15. Grinker, R. R. Neurology. Baltimore: Thomas, 1934. Pp. ix+979.
16. Kramer, F., & Pollnow, H. A hyperkinetic disorder of childhood. *Monatssch. f. Psychiat. u. Neurol.*, 1932, 82, 1. (Cited from abstract in *Arch. Neur. & Psychiat.*, 1934, 32, 857-858.)

17. LENNOX, W. G., & COBB, S. Epilepsy. *Medicine,* 1928, 7, 105-290.

18. MALZBERG, B. The prevalence of epilepsy in the United States, with special reference to children and adolescents. *Psychiat. Quar.,* 1932, 6, 97-106.

19. MCANULTY, E. A. Distribution of intelligence in the Los Angeles elementary schools. *Educ. Res. Bull., Los Angeles City Schools,* 1929, 8, 6-8.

20. NEW YORK STATE. Thirty-second annual report of the managers and officers of the Craig Colony. Sonyea: Craig Colony Press, 1926. Pp. 51.

21. ———. Forty-third annual report of the Department of Mental Hygiene, July 1, 1930 to June 30, 1931. New York: Burland, 1932. Pp. 357.

22. ———. Forty-fourth annual report of the Department of Mental Hygiene, July 1, 1931 to June 30, 1932. Albany: Lyon, 1933. Pp. 357.

23. ———. Forty-fifth annual report of the Department of Mental Hygiene, July 1, 1932 to June 30, 1933. Albany: Lyon, 1934. Pp. 357.

24. NINDE, F. W. The application of the auditory memory span test to two thousand institutional epileptics: a study in relative associability. Westchester, Pa.: Temple Press, 1927. Pp. 36.

25. NOTKIN, J. Is there an epileptic personality make-up? *Arch. Neur. & Psychiat.,* 1928, 20, 799-803.

26. PASKIND, H. A. Extramural patients with epilepsy with special reference to the frequent absence of deterioration. *Arch. Neur. & Psychiat.,* 1932, 28, 370-385.

27. PATTERSON, H. A. Syphilis and epilepsy. *Psychiat. Quar.,* 1928, 2, 63-70.

28. PATTERSON, H. A., & FONNER, D. Some observations on the intelligence quotient in epileptics. *Psychiat. Quar.,* 1928, 2, 542-548.

29. PENFIELD, W., & GAGE, L. Cerebral localization of epileptic manifestations. *Arch. Neur. & Psychiat.,* 1933, 30, 709-727.

30. POLLOCK, H. M. A statistical review of convulsive disorders in the United States. *Amer. J. Psychiat.,* 1930-31, 10, 655-661.

31. ROSANOFF, A. J., HANDY, L. M., & ROSANOFF, I. A. Etiology of epilepsy with special reference to its occurrence in twins. *Arch. Neur. & Psychiat.*, 1934, 31, 1165-1193.

32. SARGENT, P. Some observations on epilepsy. *Brain*, 1921, 44, 312-328.

33. SHANAHAN, W. T. Convulsions in infancy and their relationship, if any, to a subsequent epilepsy. *Psychiat. Quar.*, 1928, 2, 27-41.

34. STEVENSON, W. E. Epilepsy and gunshot wounds of the head. *Brain*, 1931, 54, 214-224.

35. SULLIVAN, E. B., & MACQUEEN, G. Intelligence of children suffering from allergy. (To be published).

36. TERMAN, L. M. The measurement of intelligence. New York: Houghton Mifflin, 1916. Pp. xviii+362.

37. ———. The intelligence of school children. Boston: Houghton Mifflin, 1919. Pp. xxvii+317.

38. WORSTER-DROUGHT, C. "Hystero-epilepsy." *Brit. J. Med. Psychol.*, 1934, 14, 50-82.

SUR L'INTELLIGENCE DES ENFANTS ÉPILEPTIQUES

(*Résumé*)

On a fait des mesures de l'intelligence et des observations de la personnalité d'un groupe non choisi de 103 enfants épileptiques admis à la clinique d'un hôpital. Les sujets ont été 63 garçons et 40 filles, dont l'âge a varié au temps du premier examen psychologique d'un an 10 mois à 14 ans 9 mois, et dont l'âge médian a été 7 ans 4 mois. Tous sauf un ont été de race blanche. On a mesuré l'intelligence au moyen de la révision Stanford du test Binet-Simon dans 92 cas; les autres 11 cas à cause d'un âge mental peu élevé ont nécessité l'emploi supplémentaire de l'échelle Kuhlmann Binet ou des Gesell "Normative Schedules of Development" ou de tous deux.

Pour tout le groupe le Quotient moyen d'Intelligence a été de 88, le médian de 92,4, le soixante-quinzième percentile de 106,4, le vingt-cinquième percentile de 60,1 et la variation de 11 à 141. Classifiés en niveaux d'intelligence, les même résultats ont cédé 19,3 pour cent Supérieur, 36,8 pour cent Moyen, 19,4 pour cent Normal borné 5,8 pour cent sur la Marge, et 18,4 pour cent Faible d'esprit. Le groupe entier a été mentalement détérioré comparé aux groupes de contrôle non choisis. Ainsi, 18,4 pour cent des enfants épileptiques ont été faibles d'esprit comparé à 1,3 pour cent des écoliers non choisis de Los Angeles. De la même manière on a trouvé une plus grande variabilité de l'intelligence dans le groupe épileptique. On n'a trouvé aucunes réponses épileptiques caractéristiques aux tests.

On a testé de nouveau quarante-quatre cas à des intervalles variant d'un mois à quatre ans 11 mois, l'intervalle moyen entre les tests étant 14 mois. Les comparaisons des médians sur le premier test et le second ont indiqué une détérioration progressive pour le groupe entier, mais non si grande que celle généralement rapportée par d'autres étudiants. On a trouvé une plus grande variation sur les nouveaux tests individuels pour le groupe épileptique que pour les groupes de contrôle non choisis. A l'égard de la personnalité, 44 des 103 cas ont été classifiés entre ceux qui possèdent des problèmes sérieux et continuels de personnalité ou de comportement ou les deux. Des autres cas, 45 ont été des non-problèmes. c'est-à-dire, normalement adjustés, quatre ont été des problèmes incertains, et dans 10 cas les données ont été trop insuffisantes pour la classification. On a pu grouper le comportement à problème comme l'hyperkinésie, des mauvais adjustements émotifs et d'autres mauvais adjustements de la personnalité, et des désordres du comportement. Quatorze des 19 épileptiques faibles d'esprit ont été des cas de problème. Les Quotients Moyens et Médians d'Intelligence du groupe de problème ont été moins élevés que ceux des cas sans problèmes. On n'a trouvé aucune relation signifiante entre l'âge, le sexe, la causation ou la sévérité des attaques et la personnalité.

SULLIVAN ET GAHAGAN

ÜBER DIE INTELLIGENZ EPILEPTISCHER KINDER

(Referat)

Messungen der Intelligenz und Beobachtungen über die Persönlichkeit wurden bei einer unausgewählten Gruppe von 103 epileptischen Kindern gemacht, die die Krankenhausklinik besuchten. Diese waren 63 Knaben und 40 Mädchen, deren Alter zur Zeit der ersten psychologischen Prüfung sich von 1 Jahr 10 Monaten bis zu 14 Jahren 9 Monaten erstrekte und deren Mittelalter 7 Jahre 4 Monate war. Alle ausser drei waren weisse Kinder. Die Intelligenz wurde durch die Stanford Revision des Binet-Simon Tests in 92 Fällen gemessen, die übrigen 11 Fälle wegen des niedrigen geistigen Alters erforderten einen Nachgebrauch der Kuhlmann-Binet Skala oder der Gesell Normverzeichnisse der Entwicklung oder alle beiden.

Für die ganze Gruppe war der durchschnittliche IQ 88, der Mittelwert 92,4, der fünfundsiebzigste Prozentteil 106,4 der fünfundzwanzigste Prozentteil 60,1, und die Rangordnung von 11 bis 141. Wenn die Ergebnisse in Niveaus der Intelligenz eingeordnet wurden, ergaben dieselben Ergebnisse 19,3 Prozent überlegen, 36,8 Prozent durchschnittlich, 19,4 Prozent stumpf, 5,8 Prozent an der Grenze, und 18,4 Prozent schwachsinnig. Die Gruppe im Ganzen war geistig entartet im Vergleich mit den unausgewählten Kontrollgruppen. Auf dieser Weise war 18,4 Prozent der epileptischen Kinder schwachsinnig im Vergleich mit 1,3 Prozent der unausgewählten Los Angeles Schulkinder. Gleicherweise gab es eine grössere Verschiedenheit der Intelligenz in der epileptischen Gruppe. Keine charakteristischen epileptischen Antworten auf die Tests stellten sich heraus.

Vierundvierzig Fälle wurden in Abständen von einem Monat bis zu vier Jahren 11 Monaten wiedergeprüft, indem der Durchschnittsabstand 14 Monate war. Die Vergleiche der Mittelwerte bei den ersten und zweiten Tests deutete auf eine forschreitende Entartung für die Gruppe als Ganzes an, aber sie war nicht so gross, wie von anderen Forschern allgemein angegeben wird. Eine grössere Verschiedenheit bei individuellen Wiederprüfungen wurde für die epileptische Gruppe als für die unausgewählte Kontrollgruppe gefunden.

In bezug auf die Persönlichkeit wurden 44 der 103 Fälle als im Besitz einer ernsthaften und fortdauerenden Persönlichkeit oder Verhaltensprobleme oder alle beiden betrachtet. Von den übrigen Fällen waren 45 ohne Problem, d.h. normal angepasst, 4 zeigten fragliche Probleme, und in 10 Fällen waren die Ergebnisse unhinreichend zur Einordnung. Das Problemverhalten könnte als Überbewegung, emotionelle und andere schlechte Persönlichkeitsanpassungen, und Verhaltensstörungen. Vierzehn der 19 schwachsinnigen Epileptiker waren Problemfälle. Die Durchschnitts- und Mittelintelligenzquotienten der Problemgruppe waren weniger als jene der Nichtproblemfälle. Kein bedeutsames Verhältnis zwischen Alter, Geschlecht, Ursache, oder Heftigkeit der Anfälle und der Persönlichkeit wurde gefunden.

SULLIVAN UND GAHAGAN

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

December, 1935
Volume XVII, No. 6

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Worcester, Massachusetts

Entered as second-class matter December 1, 1926, at the post-office at Worcester, Mass., under Act of March 3, 1879

$7.00 per volume
Single numbers $2.00

BI-MONTHLY

December, 1935
Volume XVII, No. 6

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

A STUDY OF THE PLAY OF CHILDREN OF PRESCHOOL AGE BY AN UNOBSERVED OBSERVER

From the Clinic of Child Development, Yale University

By

DURA-LOUISE COCKRELL

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

WLEDGMENTS

I am indebted to many people in devious ways for indirect assistance in this study, but I wish particularly to express my gratitude to Dr. Clyde M. Hill and Dr. Arnold L. Gesell for the opportunity to make this experiment; to the members of my committee, Dr. Bessie Lee Gambrill, Dr. Ruth W. Washburn and Dr. J. Warren Tilton, for the assistance, advice, and encouragement so generously given; to Miss Ann Jennings for her helpful participation throughout the experiment; and to the children and parents whose friendly cooperation made this study possible.

I wish also to express my appreciation to Helen Scott Mann, New York University, for assistance in checking and tabulating the material and to Dr. Minnie L. Steckel of Alabama College for the criticism and encouragement which led to completion.

DURA-LOUISE COCKRELL

University of Missouri
Columbia, Missouri

CONTENTS

I

INTRODUCTION

The little child's play is extremely important to his physical and mental well-being; and provision for the right kind of play is one of the major duties of those caring for young children. There is a very generally accepted and practiced theory (1, p. 120) among leaders in the field of preschool education as to the way this duty should be met. First there is agreement that adults should provide the right environment for play. Second they should remain in the background except when supervision is necessary (15, pp. 10, 45; 7, pp. 73, 90; 5, p. 84). This seems to place much responsibility upon the children and upon play materials.

Some recent studies and observations of young children offer justification for leaving the child to his own resources in this matter of play—or doing—or learning. Apparently the little child can be relied upon to react (15, pp. 2, 7), to behave in a patterned way (10, p. 302), to experiment with the attempt to master (4, p. 86), and to mature at his own rate (11). But what do playthings, of themselves, mean to children? And what happens when good materials and young children are put together and left alone? These are questions which face the mother and the teacher as they assume the duty of providing a suitable environment and direct their efforts toward remaining in the background and determining whether leaving the child to his own devices or giving him the benefit of

superior strength and wisdom will be most helpful to him in the long run.

In order to contribute further to the understanding of the play needs of young children this experiment was planned to show: (1) whether or not a change in play environment influenced the reactions of the children studied, (2) if so what differences there were in the reactions to the various materials presented, (3) the extent to which these children could or should be left to their own devices; and (4) to find any additional information concerning general or individual characteristics of play reactions as demonstrated by the children observed.

II

PLAN OF THE EXPERIMENT

A. The Procedure

The questions just enumerated are most difficult to answer because we so seldom see what a child really does alone. Parents, perhaps, have more opportunity to observe unobserved than teachers or students, but their records are rare. Although records and studies (14) have been made of the "free play" of many nursery school children, there has always been at least one teacher or one observer present and obvious. Even though these have been "in the background," it is certain that their presence was influential. No adult, no matter how withdrawn he may attempt to appear, can dispel the prestige that even a two-year old has come to associate with a grown up, nor can he prevent the young child's attempts either to gain or avoid attention. While young children adapt themselves to a considerable degree to adult observation, it is also true that they are aware of it. It is not known how stimulating or inhibiting it is, for it has been impossible to be with and observe a child who is alone.

However this impossibility has been overcome by the one-way vision screen, first put to a practical test in the Clinic of Child Development at Yale University. Briefly, "The screen functions both as a sieve and as a valve. The light, being reflected back by the screen, gives the infant no perception of the observer, while the perception of the observer penetrates the mesh. In summary, the screen allows a filtration of

light which permits him to observe without being himself observed. A mantle of invisibility, after all" (9, p. 36).

With this device available, it was possible to eliminate completely the influence of the adult and to plan a study of children's spontaneous reactions to various play materials in order to see what values these materials had *of themselves* for the children and to gain some further light upon the theory of "providing a suitable environment and leaving the children as much as possible to their own devices."

The arrangement of the Guidance Nursery at the Clinic of Child Development proved well adapted to such an experiment. Adjoining the reception room through which the children passed daily to the Nursery is a small playroom with an observer's alcove and one-way vision screen. It seems most natural and informal to invite children, at an opportune break in their play, to come into the other playroom. It was very convenient to go over familiar ground to the room which was far enough removed from the nursery to be a distinct place beyond the reach of sounds from the large playroom, yet so accessible as to offer quickly a change in play environment which was constant except for the variation of play materials presented. Here it was possible to observe without influencing and to seek the answer to the question, "Does the mere changing of available toys, the arrangement of environment, influence children's behavior?"

There were, at the time of this study, six children in regular daily attendance at the Guidance Nursery.

The backgrounds and brief pasts of these six children were well known, and they were chosen to form the permanent nursery group because of their normal physical and mental development and their good emotional adjustment.

The experiment consisted in seeing each one of these children first with every other child in the nursery group and then alone in six different play settings. Four of the settings consisted of only one type of play material, one setting offered only a bare room, a table and two chairs, while the final setting was a combination of the four different play materials. Since the children were entirely free to leave whenever they signified the desire, these play materials were each tested as to how long they could hold the children, how great a variety of activities were performed by the children with them, and how socially acceptable these activities were. Each material was seen under the same conditions in the same room with the same children, and the records were justly comparable. The actual observations in the experiment covered a period of four months and six days from November 25, 1930 to April 1, 1931. The various materials were presented alternately over this period so that the children did not know what they would find in the playroom. This scattering of the presentation of materials also eliminated any possibility of the increasing ages of the children influencing their reactions to one material more definitely than it did to the others.

In order to obtain usable information the children were seen under controlled conditions only when they

were physically well and emotionally adjusted. Since the children were allowed to attend the Guidance Nursery only when they were physically well, and since they were examined upon arrival each day, this first condition was easily met.

The second condition required careful consideration of the children and tactful approach. However, in practice it, too, was easily met, for the children were well adjusted and happy in the program provided for them in the nursery. They considered it a privilege to be invited into "the other room to play." In fact it was a favor often requested. Many times the examiner had to explain "Connie and David are coming now, you may come another time." At other times, a nod of greeting from the examiner was misinterpreted and a third child would be running toward the door as well as the chosen pair.

The very consistent cooperation of the Guidance Worker was of prime importance in maintaining this attitude in the children. She always presented the possibility of going to the observation room as a delightful one, and also cooperated in many ways in helping the examiner take advantage of natural breaks in the child's activity. The children were taken from no unusually engrossing occupation or necessary routine procedure. They were asked to leave the play room only at a natural break in their activities. The best time, of course, was to ask them to come and play immediately after their arrival in the morning. Another satisfactory time was immediately after the morning rest period. Often they were seen in the midst of the morn-

ing play period. In this latter case the examiner waited until they had finished their activity, or suggested that when they were through they might go and play. This suggestion resulted almost always in a very swift "finishing." There were two instances of the use of coercion, both occurring the first time the children were asked to "play." In one case Connie hung back. She was carried within sight of the playroom door by the teacher, who explained where she was to go and put her down to run willingly to the door. This willingness persisted throughout the experiment. At another time, David entered the playroom, not because he wanted to, but because his teacher requested it.

The verbal formula was a very simple and natural one. The first few times the Guidance Worker explained to the chosen children that there were some things to play with in the other room and they might go to them. The examiner led the way to the door, entering into any natural and cheerful conversation that arose. As the door of the playroom was opened, she said, "Here are some good things to play with. You may play as long as you want to. Come to the door and call me when you are through." A few instances of hesitancy brought the reassurance, "You can have a good time," or "Just call me. I will hear you." In a short time the children were thoroughly familiar with the proceedings and the formula was naturally simplified.

In planning the procedure, several situations were listed, any one of which concluded the observation.

1. Children calling (second time)
2. *Physical danger*
3. Danger of an emotional upset
4. Destructive activity
5. Inactivity

But in practice the conclusion of an observation was much simpler. The problem of inactivity never presented itself. There was no destructive activity. Twice the examiner stopped the observations because there was possible physical danger. Thirteen out of 216 observations were stopped because of an emotional upset, crying or quarreling or both. Twelve of the observations were stopped because of routine matters, toilet needs, need of handkerchief, or lunch time. These latter observations lasted for an unusually long time. The remaining 189 observations were definitely and calmly stopped by the children calling the examiner or going to the door and announcing "All through." Often a child called and then continued to play. The first call was ignored, but the second call was invariably answered if both children wished to leave. Usually when one child suggested going, the other was ready also, but in several cases a child succeeded in getting his companion to play a while longer. In these 189 cases, the observation did not cease until the motion to adjourn was unanimously accepted. The length of observation was dependent upon the children's ability to keep themselves happily occupied with the material.

B. The Children

The six children who served as subjects for this study were highly selected. However they represent a predominate type in many American nursery schools. Since the children in this group were chosen because of their normal physical, mental, and emotional development and because of the cooperativeness of their parents, they made excellent material for child study. This was especially so because many records of their development and of their experiences in nursery school and at home were available.

Each child had been given one or more developmental examinations at the Yale Clinic of Child Development before this study began. By the time the observations were finished, each child had been given another examination. These examinations were supplemented by interviews with parents. Records were kept of the children's adjustment, reactions, activities, and interests in the nursery. These records included the previous year for three of the children. In addition the experimenter saw every child at play in his own home and detailed records were made of their play opportunities and equipment. The parents were exceptionally cooperative in giving reports of play at home. They were also interested in the experimental procedure. Eight of the parents were observers, at one time or another, during the experiment and the incidents related after an observation or after the children were brought to the nursery proved very valuable.

The home visits and records showed the ranking of

the homes of these children to be high according to the extended Whittier scale used by Freeman, Holzinger, and Mitchell (8, pp. 23-233). Three of the homes had an index of 26, two of 25, and one of 24. The highest index obtainable by this score is 27. This was not reached because none of the children came from a luxurious home. In every case the home conditions and provisions for the child were superior. The latter included outdoor play space for all of the children, separate play rooms for four of the children, a shared nursery for one, and a play corner for another. The fathers of these children were all of professional rank. All of the mothers were home makers, though two of them were also teaching and two were in secretarial work. Each child was the first born in his family, and five of the six were only children. The family history of every child is rated as superior in the records of the Clinic of Child Development.

Medical examinations showed all of the children to be normal physically. The developmental histories, as reported by the mothers, were normal for three children and advanced for three. In the developmental examination two children were rated as high average, suggesting superiority. Two others were rated as high average, suggesting superiority, in the earliest examinations with superiority definitely established in the later examinations. Two children were rated as superior. Every child was considered well adjusted emotionally. There were, of course, outstanding personal differences but the records from the homes and the Guidance Nursery showed the children presenting

and overcoming the problems normal to their various stages of development, and responding happily and with well-rounded growth to the provisions for their care.

In family histories, native endowment and environmental advantages, the children were highly comparable. The greatest deviation in the group was in the age factor. But the experiment included an age range of particular interest to nursery schools; from just past two to almost four. (In order to assist the reader the children have been given names which correspond alphabetically to their order by age.)

TABLE 1
AGE RANGE OF THE CHILDREN IN THE EXPERIMENT

	Age at the first observation			Age at the last observation		
Child's name	Years	Months	Days	Years	Months	Days
Andrew	3	6	10	3	10	16
Barbara	2	11		3	3	6
Connie	2	10	9	3	2	15
David	2	6		2	10	6
Edward	2	4	8	2	8	14
Frank	2	2	2	2	6	8

C. THE MATERIALS

The materials used were chosen because they seemed to be often provided at home and in schools, for the indoor play of young children. They were also selected because a representative of each type of toy was possessed by every child observed in the experiment. While two of the boys owned no dolls they possessed teddy bears and doll dishes. Two of the children had a good supply of modeling clay, one had possessed

plasticine and all of the children, except one possessing clay, had been given dough to play with. Blocks and books were owned by all. The three children who had attended the Guidance Nursery the previous year had had extended experience with all of these play materials and all of the children had access to them throughout the time of the observations. Because of this it was felt that the possible effects of novelty were eliminated.

The materials presented were: housekeeping materials, blocks, pictures and books, clay and crayons, companions (no play material in the room), combined materials. While the selection of these toys was based solely upon judgment it agrees interestingly with the report of a survey made by Skalet of the play equipment in the homes of children between the ages of two and four years which appeared after the observations for this study had been made (12, pp. 258-260). From a list of 21 kinds of toys, dolls, doll dishes, and doll furniture was the type most frequently found; picture books were second; blocks, including boards and boxes were third; and handwork materials, crayons etc., were ninth.

The arrangement of the materials in the observation room had two aims which contained possible elements of conflict. It was necessary so to simplify the setting that no material (except in the last situation) would be overshadowed by another, and the children would be forced to use a certain type of material or none at all. Also, it was necessary to provide a setting which was of sufficient interest and offered enough

possibilities to the children to make the trip from the Nursery worth while and worth repeating.

The choice and arrangement of materials was necessarily arbitrary. An attempt was made to make the room attractive and conducive to activity. Three conditions were always met. First, the room was entirely cleared except for the material to be presented. A low table and two chairs were constantly present. Second, there was enough of every material for two children. There was not only enough of each material but the divisions of material were obvious and usually identical. This not only gave each child a chance to use any material as much as he wished, but it lessened chances of friction. The third condition constantly met was that each group of materials was arranged in the observation room in the same way for every combination of children or for a child alone.

Description of Housekeeping Materials. A set of unbreakable doll dishes (two cups, two saucers, one tea pot, a sugar bowl, and a cream pitcher); two doll carriages (identical, metal carriages, with rubber-tired wheels); two doll blankets, two unbreakable baby dolls; one small chest containing one doll blanket, a doll dress, skirt, and bloomers.

The table was placed against the center of the north wall so that it was immediately visible when the door was opened. On the table were the dishes arranged with a cup and saucer at each end, the tea pot, cream pitcher and sugar bowl in the center. The small chairs were drawn up beside the table and a doll was seated in the chair which faced the door. In the northeast

corner of the room was one carriage with a folded blanket hanging over the side. The other carriage, with a blanket arranged in the same way, was in the opposite corner. The chest was against the south wall. The large blanket and extra doll clothes were folded and put inside the chest. The second doll was seated on it.

In this situation, as well as the others, an attempt was made to have the room look interesting from the door, to have the materials easily available and to have the materials concentrated in the east end of the room so that the children would always be within view of the observer and yet would not be tempted to experiment with the screen. The table and chairs were always so arranged that if used in that position, the children would never have their backs to the screen.

Description of Block Setting. Three different types of blocks were arranged for the children to play with: Six hollow floor blocks, 12x10x6; smaller solid floor blocks. (These were yellow, blue, red, green, orange, and purple in various shapes, some 7x3½x1½, others 3½x3x1½; two types of cylinders, 3½ inches long or 7 inches long; and three larger colored blocks 15x7x1½.) These were all placed in a large wooden box having castors and an opening in two sides for handles. The three large colored blocks were placed on the floor beside the box; two small boxes of the Embossing Company's cubes, sixteen in a box; three unpainted planks, 48x7½x¾.

The room was arranged with the largest, hollow floor blocks piled irregularly against the north wall.

One wide plank was slanted from a block. The other planks were in two piles on the floor, one pile consisting of the two remaining wide planks, and the other of the three narrow planks. The big box of colored blocks was across the southeast corner of the room. The three largest blocks were on the floor beside the box. The table was placed with a long side against the south wall. It was near the door. The two chairs were at the west and north sides of the table and the two small boxes of blocks were placed before the chairs. Two blocks were removed from each box and placed in a tower at its right-hand side.

Description of Books and Pictures Setting. Materials: two large linen books: "A.B.C. Book," "Trains and Ships"; two small paper books: Lean's "Nonsense Book," "The Golden Goose"; two identical boxes of medium size containing pictures cut from magazines and mounted on cloth; one small decorated box of very small pictures mounted on cloth; one large box of much larger pictures, unmounted; the chest.

This time the table was placed in the center of the room. A chair was at each end. On the table in front of each chair was a linen book. The two medium sized boxes were placed one on the other, irregularly, so that there were obviously two. The two small books were also in the center of the table. The boxes and books were so arranged that there was ample room for turning pages of the large linen books. The chest was opposite the door against the north wall. The small box of tiny pictures was on it, as well as two large pictures.

On the floor near by was placed the largest box of pictures. The chest was empty.

Description of Clay and Crayon Setting. Materials: Two clay boards; two irregular lumps of clay, about one pound in weight; two boxes of Jumbo octagonal crayons; two piles of manila paper. The table was cross-wise in the center of the room, so that a child might sit at either side or end and not have his back to the screen. A lump of clay on a clay board was placed on each side of the table. A pile of paper and a box of crayons, with the lid off so that the crayons were in full view, were placed at each end. The chairs were at opposite corners of the table and could be pulled as easily to the clay as to the crayons.

Description of "Companion" Setting. Materials: the tables; two small chairs. This setting was bare and simple. The table was near the north wall. One chair was at the end of the table and one at the side. The table was not pushed flat against the wall. No other material of any sort was in the room.

Description of Combined Materials. Materials: low open shelves; one set of dishes; two books, one large and one small; two boxes of pictures, one medium, one small; one clay board, a large lump of clay; a pile of paper; a box of Embossing Company's blocks, large size, 36 in a box; two dolls; one carriage and blanket; the chest and extra doll clothes; six floor blocks. In this setting each of the previous situations was represented. The shelves were against the north wall. The set of dishes was arranged on the top shelf, with the tea pot, sugar bowl, and cream pitcher in the

center, and a cup and saucer at each side. The next shelf contained the clay board and clay, two books and the boxes of pictures. On the bottom shelf were the box of blocks, a pile of paper and a box of crayons.

The table was lengthwise across the room, with a chair at each end. In the northeast corner was the irregular pile of floor blocks. The doll carriage was in the southeast corner of the room, and the chest containing the clothes, with the two dolls seated on top, was near by. This set-up was similar to many home playrooms for young children, thoughtfully but not elaborately furnished.

D. THE RECORDING

In order to establish methods of procedure and recording, 17 observations had been made of children alone in the Guidance Nursery of the Clinic of Child Development. Eleven different children were seen in various combinations of five, three, two, and one. Nine different materials were presented to the children. This preliminary study could only offer indications, but they were definite and interesting enough to provoke further investigation. The tangible results of these observations were a system of recording, a selection of materials, and the plan of further procedure.

The method of recording grew from a diary record form which presented the whole picture of the children's reactions, physical, verbal, and emotional, while they were under observation. This whole picture was extremely desirable, but the method involved and re-

quired a great deal of interpretation when records were compared. However, one very objective and meaningful piece of information which could be drawn from the diary record was the number of different uses found for the material and the variety of activities performed by the children. Another very objective and comparable part of the data was the time element. The number of minutes children voluntarily stayed with various toys was a measure of the interest the toys held for them. However, merely noting the time spent in the room was not a complete indication of the "holding power" of the material, for the children spent varying amounts of time in attending to one another, in running about and other innumerable activities in no way related to the material at hand. These activities, when closely observed, were matters of seconds, not minutes.

This was brought out over and over. For instance, a diary record would note, "Child enters the play room, sits on the low bench, kicks his heels against the bench-side and looks around the room. Then he runs to the floor blocks and piles one upon another, saying 'one here—one here'. The blocks fall."

When this activity was recorded with a stop watch, with numbers indicating minutes and seconds from time of entering the room, the record became:

Enters room	0- 00
Sits on bench	0 -10
Kicks heels	
Looks about	

To floor blocks	0 -30
Piles blocks	
Talks	
Adds third block	1 -50
Talks	
Adds 4th block	2 -10
Adds 5th block	
Adds 6th blocks	
Blocks fall	2 -45

In the diary record the twenty seconds of heel kicking vied in importance with the two minutes and fifteen seconds of constructive block play. Even to record this to the nearest minute would skew the picture decidedly, for the bench-sitting and heel kicking would appear to be half as long as the block play when in reality it was less than one-sixth as long. Therefore the practice became to enter time to the nearest five seconds.

The diary records included much language. This was very enlightening but often so constant that recording it sacrificed other activities and details more important. It was decided to make no attempt to record language in any completeness, but to list "talks" as an activity, just as "sings," "cries," or "runs" was listed. Actually, many remarks were recorded verbatim when it was practicable and they were pertinent.

The recording of time and the listing of activities in a running account gives very definite quantitative data by which to judge the value of materials for children and the children's ability to use materials. Since

the activities could be classified as socially acceptable or non-acceptable, a still further criterion was added as a basis for judgment.

This type of record, with the entry of time in small units, was impossible to keep for several children using a variety of materials. But the arrangement of only one type of material in an otherwise bare room made that material the center of attention, and the use of it by the children comparable to their use of some other material presented in the same way, and had the additional value of simplifying the recording.

It was found to be possible to record very fully the behavior of two children, but not more than two, in such a controlled situation. The method adopted finally was that of recording observations in parallel columns for the two children observed. The items noted were (1) material approached (approached was interpreted as meaning attended to and within reach), (2) the time of approach (to the nearest five seconds); (3) shifts in activity while with the material, and (4) the time of the shift in activity (Record No. 75).

After the final method was decided upon, five observations were recorded to give the examiner further practice in the use of the method. Two of the observations were on one type of material, "housekeeping materials," and three of the observations were of "combined materials," a more difficult but a possible set-up to record. The use and comparability of such records were further tested in the fact that four of the five observations consisted in seeing one of a pair of identical twins with another child in a play situation, and

then seeing the other twin in the same situation with the same child. The resulting records not only presented a clear, whole picture of each twin's reactions, but they were quantitatively comparable.

Reliability of the Records. Since every observation included in this study was recorded by the experimenter, it is safe to assume that each record is comparable to any other and that the percentage of error is constant. But to test further the reliability of the method and the ability of the experimenter to record, simultaneous records were taken. There was very definite agreement in time and in number of activities. The records used as a check were made by a member of the Clinic staff. One time the method was for the experimenter to take the record as usual while the second recorder made a check mark for each activity—that is, shift in activity. When the recorded activities were counted there were 36 for Barbara, and 27 for David. The number of activities checked was 36 for Barbara and 29 for David, indicating that the entry of time and writing of records did not cause the overlooking of an appreciable number of activity shifts.

Another record covering 23 minutes gave 58 time entries for Edward by the observer and 61 entries by the second recorder. The observer had 60 time entries for Connie while the second recorder had 58. Of these time entries the observer had 20 designating definite shifts from one material to another, while the second recorder had 21 for Edward. All of the recorded shifts agreed as to material approached, 13 agreed exactly as to time, while 6 differed by five seconds and

one differed by ten seconds. For Connie there was even closer agreement in shifts. Twenty-eight shifts agreeing as to material approached were recorded by both observers. Eighteen of these shifts agreed to the second, seven differed by five seconds, and three differed by ten seconds. Of the totals of seconds spent with material, taken from these simultaneous records, 42 agreed, 28 were five seconds off, and six differed by 10 seconds.

The several instances of five seconds difference seemed due to the fact that when the hands of the stopwatches were approximately midway between two numerals one recorder noted the number ahead, while the other entered the numeral just passed.

The general picture of the type of play, the activities with each material, the material holding the child a long time, and the material holding him but a short time agreed consistently and gave further evidence that the method of recording was valid and usable.

Record No. 75 Record No. 75

Clay and Crayons

Andrew *Barbara*

Both at the chairs 0-00
Pull chairs to the clay and sit by the table
(The sun is shining on the table.)

"Sunshine."

Both pull pieces from the lump of clay.

"I goin' make an Easter rabbit." 0-15

"I make his shoes."	"I make a shuah rabbit." (looks questioningly)
Rolls some clay	"Didn't I?" 0-35
	Also rolling clay 0-40
"Mine's going to be a big one."	"Mine gonna be a big one."
	Laughs 1-0-0
Andrew sticks his roll on the lump 1-15	Still rolling clay
"Look, I makin' a Easter rabbit."	Pinches some off 1-25
Rolls a piece again	
"Roll it this way."	
Rolls swiftly	
"Do that."	Is also rolling 1-40
"I'm making an Easter rabbit." 1-50	"I are too."
Sticks two pieces firmly together 2-00	Watches the way Andrew works
Pinches at the pieces 2-15	
Molds them with his fingers	"I have something."
	Holds a piece toward Andrew 2-25
Andrew continues with his clay	"I make a rabbit."
Pats the clay flat 2-45	
Rolls it between his hands 2-50	Also pats 2-50
Continues rolling the clay	Pats her clay again 3-00
Still rolling 3-10	
Presses it onto his lump	Holds a piece toward Andrew 3-20
"I'm through." 3-30	"I'm not."
	Drops a piece and gets it 3-35
	Rolls it

"When I get this done."			
He rolls a large piece about two inches long.			
"Look what I makin'."	3-55		
Sticks it on the lump.			
"I made a candle."	4-10	"I makin' candle."	4-15
Takes a larger piece of clay		Rolls clay	
Makes a bigger roll			
Calls once		"I not goin' out."	4-35
Continues with the clay		To the crayons	4-40
Pats the clay			
Also pulls a box near	4-50	Sticks a crayola in her clay	
Takes the white crayon	4-55		
Marks on the board			
Marks on the lid			
Marks on the paper		Marks on the lid	5-05
"Lookit."		Singing	
Marks on the box lid	5-25	Marks inside the lid	
Calls	5-30	"Don't call Miss Cockrell." Marks	
Puts the crayon in the box			
"I goin' out."	5-45		
To the door		Also to the door	5-50
Calls		Calls	

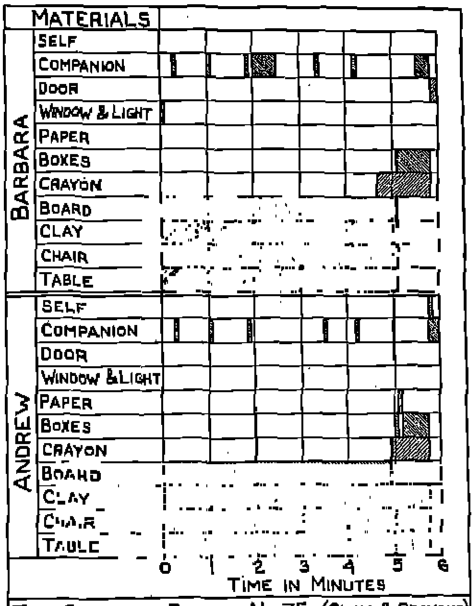
(Observation ceases 6-00)

E. The Tabulation

The records taken in this way contained definite, full information concerning the behavior of the chil-

dren in a certain setting, but this information was not observable at a glance and the records were not immediately comparable. Therefore methods of tabulation were devised. The running account of each observation was recorded in two organized forms; the first a time graph, and the second an activity chart.

The record of an observation, following as it did the swiftly changing activity of the children, accounted for the way time was spent during the observation, but it was difficult to ascertain quickly the total amount of time spent with any one object in the room; and the mere addition of seconds was apt to become confused because often a child would be attending to or using



TIME GRAPH FOR RECORD NO. 75 (CLAY & CRAYONS)
FIG.-1

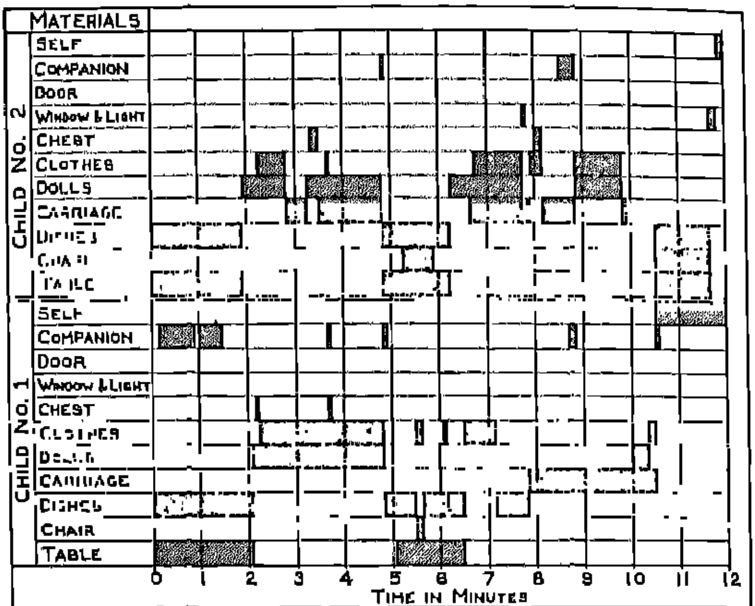


FIG. 2 - TIME GRAPH FOR HOUSEKEEPING MATERIALS

more than one thing at a time. A need was felt for a graphic way of presenting the division of time which would preserve the picture of quickly shifting activity and also make clear the inclusiveness as well as the length of the interest span.

The time graph fills this need. Examples are given in Figures 1 and 2. The graphs for a single observation consisted of two parallel graphs, where two children were observed, one for each child, corresponding to the parallel columns in the records. The numbers running horizontally on the graph indicate minutes. Each square represents 60 seconds and time can be graphed to indicate the five seconds recorded in the ob-

servation. Vertically there is listed every object in the room including "self" and "companion." By glancing along the row representing any object one can quickly discover the total number of seconds spent with the object and see how long or how short the attention spans are. Vertically on the graph it is easy to find what the children are attending to during any given five seconds and also how many or how few things were attended to at any one time. Since the records included the time entries the graphs were readily compiled. Some questions, however, were presented. Very occasionally a shift in activity would be recorded without a time entry. These were obviously brief ones such as "answers companion" or "looks toward window." It was decided to graph any activity recorded without a time entry as being five seconds in duration since no smaller unit of time was recorded and the attention span was evidently so short as to make the recording of time difficult.

Shifts of attention from material to material were easily recorded. They were marked by the children's overt acts of touching a material or looking at it, or both. But attention to companion was not always so clearly determined and required more defining.[1] Sometimes the attention was indicated verbally, sometimes by a gesture. After considering the many ways of noticing the companion which were included in numerous records the following were adopted:

[1]Other investigators have met difficulty in defining and timing social responses, i.e., Bott (2, p. 81).

1. Actual physical contact with companion
2. Exchange of objects with companion
3. Simultaneous use of an object with companion
4. Response to companion (either verbally or by change of activity)
5. Attempt to change companion's activity
6. Watching companion

Any one of these criteria was considered sufficient basis for recording. The first four always involved both children observed. The last two might involve only one child. The fifth criterion is somewhat dependent upon interpretation. Response to companion is definite and follows a successful attempt to change another's activity. But besides these successful efforts in directing the behavior of another, there were fruitless attempts to command or instances of "bidding for attention" which were so definite as to necessitate consideration. Since two- and three-year olds are not particularly subtle in attempts to direct or gain attention and since the one examiner made all interpretation it is felt to be consistent.

Play alone, graphed as "self," was that play which concerned neither the material in the room nor the companion. Such activities as trotting around the room empty-handed, running and sliding, sitting on the floor, and chanting were considered as "self-play".

With these definitions established and the graphs made, an analysis of the distribution of time under various categories was made possible. A significant division of time, both to psychologists and educators,

is the attention span. Most of the studies of the play of preschool years have considered this of major importance and two recent studies of preschool children have dwelt with it exclusively (6, 13). But again a definition was needed.

Although the study of the interest span of preschool children by Herring and Koch appeared after much of the experimentation and graphing in this investigation was completed, it was found that the time graphs indicated interest spans as defined by these authors and their definition was adopted because of its suitability and in order to facilitate comparison of results. Their explanation is:

> "It is probably clear that from our point of view one activity was differentiated from another primarily on the basis of the toy which centered it, i.e., a child was considered to have shifted his interest when he turned from one toy to another. Sometimes, however, two toys were employed simultaneously . . . . Under these circumstances when a shift was made from the activity combining two toys to an activity involving only one of the toys and that one not the toy around which most of the previous play had seemed to focus, such a shift was recorded as initiating a new activity. Thus if, after hauling the acorns to an imaginary town in the truck, the child suddenly attempted to throw acorns through a transom we took the stand that he had engaged in two activities instead of one (13, p. 253).

In this definition shift in interest and shift in activity are considered to be the same. In the present study activity is given a more limited meaning which will be fully discussed. But, since a child was considered to have shifted his interest when he turned from one toy

Activity Chart

Clay and Crayons

Barbara (and Andrew) No. 75

TABLE	CHAIR	CLAY	BOARD	CRAYONS	BOXES	PAPER	WINDOW & LIGHT	DOOR	COMPANION	SELF
Sits	pulls	pulls	clay	from box	removes	talks		goes	talks to	calls
beside	sits	"Shoe	on	in clay	crayons		looks	to	questions	
	in	rabbit"		marks	marks		toward		watches	
		rolls		lid	lid		"sunshine"		talks to	
		laughs		sings					shows	
		pinches		marks					clay	
		rolls							talks to	
		talks							shows	
		shows to C.							clay	
		talks							inhibits	
		pats								
		shows to C?								
		"Rabbit"								
		drops								
		retrieves								
		rolls								
		"candle"								
		crayon in clay								

Activity Chart

Clay and Crayons

Andrew (and Barbara) No. 75

TABLE	CHAIR	CLAY	BOARD	CRAYONS	BOXES	PAPER	WINDOW & LIGHT	DOOR	COMPANION	SELF
Sits	pulls	pulls	clay	from box	removes	marks		goes	talks to	calls
beside	sits	"Easter	on	marks	crayon	on		to	talks to	calls
	in	rabbit"	marks	board	replaces				talks to	calls
		rolls	on	marks	crayon				answers	
		talks		lid					answers	
		on lump		marks					answers	
		talks		paper						
		rolls		talks						
		2 pieces		marks						
		together		lid						
		pinches		in box						
		molds								
		pats								
		rolls								
		presses								
		talks								
		rolls								
		on lump								
		"candle"								
		pulls								
		rolls								
		pats								

to another, the method of graphing records the interest span as Herring and Koch defined it.

The activity charts follow a simpler plan. A chart was made for each observation of each child. Again every object in the room was listed including "self" and "companion." Under each object were tabulated the activities the child carried on in connection with that object. The same six criteria listed for the time graphs (p. 27) were used to determine what activities to list under companion. But again there was the difficulty of tabulating an activity which involved both objects and the companion such as "pretending to pour tea for the companion." In such cases "pours for C" was listed under "dishes" and under "companion" and was ringed under "dishes" so that one activity would not be counted twice. From these charts could be found the total number of activities connected with one object, the number of activities performed by one child during one observation, the variety of activities, and the repetition of activities.

It was necessary, in making these charts, to define activity. The children presented a picture of continuous activity, which was ever-changing and had definite shifts, but never a clear-cut break. The records noted shifts from one material to another or shifts in type of activity with a single material. The time graphs presented the shifts from material to material. The activity chart listed every change in activity recorded, using the word descriptive of that change and counting each word as "an activity" because the phrase is simpler and more usable than "A shift in activity."

Neither the records nor the activity chart attempted to consider the many motions made by the children. Since "activity" is to be used continuously throughout this study, an example of its meaning will be helpful. In Record No. 75 the observation made for Andrew between 1-50 and 3-30 is tabulated as:

Clay
2 pieces together
Pinches
Molds
Pats
Rolls
Presses
Talks

In the record it will be seen that "rolls" is recorded three times, but it is called only one "activity" under "clay" in the chart, for, though it continued long enough to be mentioned three times, there was no change in the activity. A sample activity chart is given (page 32). Here it will be seen that "pats" is counted "one activity," "rolls," another, but more accurately they are different expressions of a continual flow of activity, the end of one marking the beginning of another.

III

COMPARATIVE VALUE OF PLAY MATERIALS

A. Children's Preference for Materials

The method of procedure and observation just described subjected each of the selected play materials to controlled conditions which were very nearly identical. The recording and tabulation preserved and organized the resulting reactions of the children so that they were quantitatively comparable and offered a basis for the evaluation of the materials. The first piece of information available was that of preference. A play material has met a severe test if it has held children in an otherwise bare room even though they were entirely free to leave at any time and return to an attractive and beloved nursery. The play materials selected for this study were put to such a test that the number of minutes the children remained in the room with each material provided an objective measure of the "holding power" of the various materials and a measure of the children's preferences (Table 2).

TABLE 2

Number of Minutes Spent With Each Material

Rank	Material	Minutes	Multiples of time for companions
I	Combined materials	786½	2.7
II	Clay and crayons	673½	2.3
III	Pictures and books	479	1.6
IV	Blocks	461½	1.5
V	Housekeeping materials	381	1.3
VI	Companions	287½	1
	All materials	3,069	

A comparison of the total number of minutes recorded for each material under the controlled conditions of the experiment shows distinct differences. To facilitate this comparison the minutes in Table 2 are also expressed as multiples of the time for Companions, showing the relative holding power of the materials. Ordinary observation would lead to the conclusion that children would willingly stay in a well-supplied play room for a longer time than they would in a room offering nothing to play with. But the other situations included materials accepted by nursery schools as good play equipment. Yet they have different values to the children as judged by this method.

In order to check the reliability of these records further the data were divided into halves as a substitute for repeating the experiment. Division was made by the odd-even method, and, as additional check, by children's ages so that figures for the three younger children were compared to those for the three older children. The complete records[2] include this double test for every table presented in this report. That the differences in time spent with each material as shown in Table 2 are reliable ones was shown when the materials maintained the same rank when divided by the odd-even method and remained constant when judged by time for the three older children or the three younger children.

It is evident that, of themselves, clay and crayons are interesting to these children for a longer time than

[2]Complete tables are in the dissertation filed in the Library of Yale University.

III

COMPARATIVE VALUE OF PLAY MATERIALS

A. Children's Preference for Materials

The method of procedure and observation just described subjected each of the selected play materials to controlled conditions which were very nearly identical. The recording and tabulation preserved and organized the resulting reactions of the children so that they were quantitatively comparable and offered a basis for the evaluation of the materials. The first piece of information available was that of preference. A play material has met a severe test if it has held children in an otherwise bare room even though they were entirely free to leave at any time and return to an attractive and beloved nursery. The play materials selected for this study were put to such a test that the number of minutes the children remained in the room with each material provided an objective measure of the "holding power" of the various materials and a measure of the children's preferences (Table 2).

TABLE 2

Number of Minutes Spent With Each Material

Rank	Material	Minutes	Multiples of time for companions
I	Combined materials	786½	2.7
II	Clay and crayons	673½	2.3
III	Pictures and books	479	1.6
IV	Blocks	461½	1.5
V	Housekeeping materials	381	1.3
VI	Companions	287½	1
	All materials	3,069	

A comparison of the total number of minutes recorded for each material under the controlled conditions of the experiment shows distinct differences. To facilitate this comparison the minutes in Table 2 are also expressed as multiples of the time for Companions, showing the relative holding power of the materials. Ordinary observation would lead to the conclusion that children would willingly stay in a well-supplied play room for a longer time than they would in a room offering nothing to play with. But the other situations included materials accepted by nursery schools as good play equipment. Yet they have different values to the children as judged by this method.

In order to check the reliability of these records further the data were divided into halves as a substitute for repeating the experiment. Division was made by the odd-even method, and, as additional check, by children's ages so that figures for the three younger children were compared to those for the three older children. The complete records[2] include this double test for every table presented in this report. That the differences in time spent with each material as shown in Table 2 are reliable ones was shown when the materials maintained the same rank when divided by the odd-even method and remained constant when judged by time for the three older children or the three younger children.

It is evident that, of themselves, clay and crayons are interesting to these children for a longer time than

[2]Complete tables are in the dissertation filed in the Library of Yale University.

books, blocks, or dolls and dishes. It is also evident that a room with many toys is a greater rival to a pleasant nursery than a room with no toys. This evaluation of play materials by the length of time they will hold children, without adult stimulation, from other interesting places is a basic and pragmatic one. It is an important one to a busy mother who decides, "Well, I found a good thing for Jane to play with. She worked a little piece of dough the whole time I was getting dinner." But this is not the only factor in the value of play materials. Other questions immediately arise. Why is there a difference between the materials when they are measured by time? What is it that makes children stay longer with one material than another? And how are these minutes spent?

B. Activities With the Materials

Total Number of Activities. The recording and tabulating of the constant activity throughout all the observations have been described. The distribution of

TABLE 3
TOTAL NUMBER OF RECORDED ACTIVITIES WITH EACH MATERIAL

Rank	Material	Number of activities
I	Combined material	4501
II	Clay and crayons	2900
III	Blocks	2092
IV	Housekeeping materials	1948
V	Pictures and books	1917
VI	Companions	1510
	All materials	14868

these many activities among the various play materials is shown in the following table (Table 3):

It might reasonably be expected that the ranking of materials according to the number of activities would be identical with that according to the number of minutes, since more time would provide for more activities and more activities would necessitate more time. But Table 2 and Table 3 do not show identical ranking. When they are compared it will be noted that Pictures and Books drop from third to fifth place and Blocks and Housekeeping Materials rise accordingly. Pictures and Books held the children 58 minutes longer than Housekeeping Materials, yet the children performed 31 more activities with the Housekeeping Materials.

When the data for recorded activities were split by the odd-even method they did not show the same consistency evidenced by the data for time for Pictures and Books interchange rank with Housekeeping Materials. The activity of the three older children ranked the materials just as Table 3 does, but Blocks took fifth place with the younger children in number of activities. The extremes are the same in Table 2 and Table 3 and remain the same in every division of data. Combined Materials is consistently first, Clay and Crayons second, and Companions last. But the other materials do not agree in rank when measured by the two methods. This partial consistency leads to the conclusion that the length of time children spend with a material is not entirely dependent upon, or related to,

the number of activities the children perform while with the material.

Rapidity in Shifts of Action. Observation of young children at play might easily lead to the assumption that the more opportunity a child has for activity the more he enjoys himself, and the more a toy permits swiftness and abundance of activity the better he likes it. From the data in Table 2 and Table 3 the rapidity of activity is found by computing the number of recorded activities per minute.

TABLE 4

NUMBER OF RECORDED ACTIVITIES PER MINUTE

Rank	Material	Number of activities per minute
I	Combined materials	5.72
II	Companions	5.28
III	Housekeeping materials	5.11
IV	Blocks	4.53
V	Clay and crayons	4.29
VI	Picture and books	4.00
	All materials	4.84

These figures (Table 4) illustrate quantitatively the impressions received during the observations of activity with the Combined Materials and the Housekeeping Materials, of restless activity in the bare room, of steady, purposeful activity with the Blocks, and of quieter play at the table with Clay and Crayons and Pictures and Books. They explain the discrepancies between the amount of time spent and the number of activities with materials.

But this ranking in no way agrees with that of Table

2. And when these data were divided a variation in swiftness of activity was shown for Companions and Pictures and Books. The shift for Companions is to be expected, for, while the children usually reacted to the bare room with swiftly changing activity, running, jumping, or climbing, there were occasions when a child would lean on the table, look about the room, remark that there were "no things" and call to leave. The general reaction to Pictures and Books, however, was more consistent. The children appeared very mature and poised as they sat down at the table arranged for them and looked at the books. They were thoroughly childish as they sat on the floor and placed the pictures about, and usually they were more quiet and "slow moving" in this situation than in any of the others. Occasionally, however, they simply ran and jumped about the room for a few moments, ignoring the play material provided. Nevertheless the data distinguished between the three set-ups in which swift activity occurred and the three wherein activity was slower. While it seems reasonable to conclude that, within the limits defined by the range of rapidity in activity itself, the swiftness in the shifts of children's activity in play is altered by the type of material they are given, yet the "holding power" of this play material is in no way related to the rapidity of activity involved in its use—and rapidity of shifts in activity in no way corresponds to the number of minutes children remained with any of the play materials considered in this study.

Repetition of Activities. The constant and rapid

activity entailed a great deal of repetition in behavior. From the activity charts it was possible to count the number of activities that were recorded more than one time. The first record of an activity in the charts of any child was considered "original" for that child. Any other records of this activity were checked as "repeated." When the materials are ranked according to the percentage of repetition involved in their use differences are disclosed (Table 5). A comparison of this

TABLE 5

PERCENTAGE OF REPETITION IN ACTIVITIES WITH MATERIALS

Rank	Materials	Percentage of repetition
I	Clay and crayons	66
II	Companions	62
III	Combined materials	60
IV	Blocks	58
V	Housekeeping materials	58
VI	Pictures and books	55
	All materials	59

table with the previous ones shows that the amount of repetition in activity is not related to the total number of activities recorded, nor to the rapidity of these activities, nor to the number of minutes spent with a material.

Division of this data by the odd-even method showed that the materials are not consistent within themselves in percentage of repetition. When the percentage of repetition in the activity of the three oldest children was compared with that of the three youngest there was again inconsistency of ranking. But consistency appeared in one matter. That is, the percentage of

repetition in the activity of the younger children was greater than in that of the older children.

These findings agree with the general observation that repetition is noticeable in the play reaction of young children. Slightly more than half of the activities performed by the children studied were repetitive ones. About the same amount of repetition, an average of 59 per cent, appeared in each situation. All of the materials had sometimes slightly more and sometimes slightly less than 59 per cent repetition, except Clay and Crayons, and Companions. These two situations ranked consistently high. It might be inferred from this that materials such as clay and crayons, requiring the same type of manipulation even when different things are made, will increase the usual amount of repetition and that a scarcity of play material, in restricting possibility for varied activity, will also increase repetition. There are indications that the amount of repetition decreases between the second and the fourth birthday. But one conclusion is evident. More than half (between 51 per cent and 71 per cent) of the unguided play activity of these children was repetition, no matter what play situation they were put into. Also the amount of repetition is not coincident with the holding power of materials.

Variety of Activities. Recognition of the existence of repetition should not conceal the equally important recognition of variety in the play reactions of these children. There seemed, to the examiner, to be an amazing amount of it. Adults, it usually appears, have grown to have rather circumscribed ideas in con-

nection with toys and their uses. For instance, they may see possibilities in a doll carriage. A little girl can roll her doll in it, and have the delights of dramatization, and imitating her elders, as well as physical activity. These children, left to play alone with a doll carriage, used it to roll dolls, but they were in no way limited by this adult conception of a carriage which was also used to roll dishes and doll clothes. It was a bed, a train, a fire engine, and a car. Besides being a thing to roll, it was a thing to push and chase, a thing with a moveable hood to tip and with screws to turn. It was rolled on its front wheels and on its back wheels, it was put on its side and two wheels were whirled, it was turned upside down so the four wheels whirled in the air. These are only a few of the various activities performed with a rather unadaptable toy.

In order to determine the amount of variety in the children's play these different activities, taken from the original activities in the activity charts, were listed for each type of material. One kind of activity was listed only once even though it had been performed by all six children, or many times by all six children. The value of this enumeration will be obvious when it is recalled that each material was presented and observed 20 times. Each material was subjected to the same conditions and the children reacted favorably to every material so that there were at least four hours and 47 minutes of observation and 1510 recorded activities. Of these activities over one-half were repeated so that ample and equal opportunity was afforded for using the materials in all the different ways possible to them and the chil-

dren. A count of these activities gives a definite, comparable measure of the variety of possibilities each material offered the children in their play (Table 6).

TABLE 6
VARIETY OF ACTIVITIES WITH THE MATERIALS

Rank	Materials	Number of different activities
I	Combined materials	829
II	Clay and crayons	496
III	Pictures and books	423
IV	Blocks	419
V	Housekeeping materials	378
VI	Companions	276
	Total	2821

This ranking agrees exactly with the ranking of materials according to the total time spent with them (Table 2), and leads to the conclusion that, while much activity and numerous repetitions were evidenced in every situation, the factor which determined the length of time a material would be used was the variety of activities to which it lent itself in the children's play. The quality of offering many possibilities has often been extolled as important in play material, especially from an educational point of view. It would appear that little children also appreciate that quality and, though for less learned reasons, they prefer playthings which permit activity leading to further activity.

Social Acceptability of Activities. These tables describing quantitatively the characteristics of the spontaneous play of some young children are interesting to a student of child behavior. But the practical adult responsible for a young child will question the impor-

tance of knowing that children played undirected for a comparatively long time with clay, and that their activity was more or less rapid, and more or less repeated and full of variety, unless it is also known how much of the activity was worth while and how much was harmful. The dependence of the amount of "goodness" or "badness" in behavior upon the criteria of judgment is evident and the difficulty of finding such criteria that are generally accepted is obvious.

When the children and materials were viewed dispassionately through the screen, the materials being regarded absolutely as things to be used in any way, the only opportunity for the children to misbehave was to harm themselves or another child. All precautions were taken in arranging the situations so as to remove possibilities of injury with materials or furniture. And the children proved to be exceedingly wise and capable in avoiding danger to themselves. They were unexpectedly concerned with the safety of their companions also. David tumbled from a pile of blocks but shed no tears. The next day he saw Edward climbing similarly and warned "Don't fall." Andrew noticed Barbara's shaky position on some blocks. "Do you want to fall? That isn't salty enough to stand on."

There was an extremely small amount of friction or personal attack between any of the children. There were only five cases of misbehaving according to the very liberal standards of the observational set-up. Three times pushing another child caused tears and stopped the observation. Two of these pushes were in anger, one was experimental. There was one tussle

which ended finally in laughter and continued play. There was one slap which was as surprising to the slapper as to the one slapped. It was followed by an awkward silence, both children hanging their heads. Finally one looked up, began to talk, and play was resumed to continue amicably for ten more minutes.

But the experimental criteria used above do not agree with those of the usual home or nursery school situation. From this viewpoint rather than the experimental one the following criteria were set up for judging socially non-acceptable activity:

1. Activities involving discomfort for the companion.
2. Injury, or threatened injury, to self. (The only cases of this, except for two slight falls, were those of putting things into the mouth.)
3. Injury, or threatened injury to materials or furniture.
4. Emotional display of anger or fear.

"Threatened injury" was included because often an act which was injurious at one time happened not to be at another time. For instance, standing on the table did not always scratch it, and continuous pushing of the light button did not "wear it out" as Andrew suggested it might. No child was hurt by what he put in his mouth. But such activities threatened injury and were considered unacceptable. With these criteria it was possible to check and count the socially non-acceptable activities in the activity charts (Table 7).

This consideration of the materials presents a rank-

TABLE 7

Percentage of Socially Non-Acceptable Activities with Materials

Rank	Material	Percentage of non-acceptable activities
I	Companions	21
II	Clay and crayons	8
III	Combined materials	7
IV	Pictures and books	5
V	Housekeeping materials	4
VI	Blocks	2
	All materials	8
	All materials (except companions)	5.2

ing unlike any of the previous ones. (Probably it should be reversed since Blocks are the "best" material, in that play with them is more socially acceptable). But each material maintained the same place when the data were divided. The significant fact brought out in this comparison of materials is that the mere changing of the type of play material, in situations which were otherwise identical, changes the amount of socially non-acceptable activity in the play of children.

There still remained some activities which were not harmful, but which might be annoying and were, from a critical standpoint, a misuse of materials. Examples of these are: stacking doll dishes in the doll carriage; inhibiting the companion; scattering the blocks after building with them; or simply yelling. When much stricter and more empirical standards are set up for judging activity, checking not only those which might be harmful, but those which were questionable or annoying, the ranking of the materials remains the same (Table 8).

This table emphasizes another important fact in the

TABLE 8

PERCENTAGE OF SOCIALLY NON-ACCEPTABLE AND ANNOYING ACTIVITIES WITH MATERIALS

Rank	Materials	Percentage of annoying activities
I	Companions	27
II	Clay and crayons	10
III	Combined materials	10
IV	Pictures and books	7
V	Housekeeping materials	7
VI	Blocks	5
	All materials	11

comparison of materials. A bare room forces the percentage of non-acceptable activities unusually high. It is about three times greater than in a well-supplied play room. When annoying behavior is checked as well as "bad" behavior, each material has 2 per cent or 3 per cent more of non-acceptable activities except Companions, which has 6 per cent more. The fact that there was less to harm in this situation did not counteract the fact that there was nothing to play with. A consideration of this effect of an empty room upon these healthy, intelligent, adjusted children should throw some light upon the effect of the type of room, in which children often find themselves, wherein there is much to harm, yet nothing with which to play.

Further characteristics of the spontaneous play of the children are brought out by these figures. The children's preference for materials apparently was in no way related to their "good" or "bad" behavior while with the material. There is no evident relationship between the rapidity of activity and its non-acceptability or between any of the characteristics tabulated,

but there is consistency within the materials themselves.

Also the amount of non-acceptable activity is surprisingly low. When the children were given absolute freedom in a room *provided* for play and then their behavior was judged by the standards of a restricted situation (an injustice which fortunately had no effect on the children), an average of only 5.2 per cent of their activity was unacceptable.

C. Distribution of Time Spent with Materials

Length of Observations. The total time of spontaneous play with a toy has been considered in this and other studies as a measure of the "holding power" of the plaything and an indication of children's preference. The records of this experiment not only offered information concerning the total time spent in the various play situations, and the characteristics of the activities involved, but also presented information on the way the time was spent during each observation of each material. Data on the length of observations, depen-

TABLE 9

Average and Range of Length of Observations

Rank (recording average length of observations)	Material	Average length of observations	Range in length of observations
I	Combined material	21 min. 48 secs.	5 min. to 41 min.
II	Clay and crayons	18 min. 36 secs.	½ min. to 33 min.
III	Pictures and books	13 min. 18 secs.	2 min. to 23 min.
IV	Blocks	12 min. 48 secs.	2 min. to 27 min.
V	Housekeeping materials	10 min. 30 secs.	1½ min. to 20 min.
VI	Companions	8 min.	½ min. to 21½ min.
	All materials	14 min. 12 secs.	½ min. to 41 min.

dent entirely upon the children's reactions, are presented in Table 9. The ranking, of course, agrees with Table 2.

The figures in this table compare interestingly with those found by Bridges (3) in a study of ten children ranging in age from two years six months to three years eight months. She studied a number of materials in a Nursery School environment and measured the time of an "occupation" defined by the child spontaneously taking a material and spontaneously returning it to the cupboard when through. She found the average length of an occupation to be eight minutes and the longest time spent in one occupation to be 37.5 minutes for boys and 34.5 for girls.

Length of Interest Span. Probably the most significant division of time within an observation is that of interest spans. The method of defining these has been discussed previously (pages 25 and 26). Table 10 presents the information concerning interest spans derived from the time graphs.

TABLE 10

LENGTH OF INTEREST SPANS WITH EACH MATERIAL

	Material	Average interest span (seconds)	Average of maximal interest span (seconds)	Longest interest span (seconds)
I	*Clay and crayons*	115.63	542.31	1230
II	Combined materials	93.67	393.23	1655
III	Blocks	89.64	309.71	800
IV	Pictures and books	64.87	203.17	450
V	Housekeeping materials	51.67	189.68	620
VI	Companions	45.53	146.77	330
	All materials	75.86	297.48	
	All materials (except companions)	81.53	327.62	

Since the definition formulated by Herring and Koch (13) was used so that their study and this one measured the same thing, the results are comparable to a certain extent. There are several differences in procedure, particularly in the fact that in the quoted study children were presented with six toys of different types at one time and that no distractions less than 30 seconds in length were recorded. The time graphs of this study included distractions as short as five seconds.

Herring and Koch found that the averages for the maximal spans, by the Thirty-Second Scheme, were 514.5 and 456.1. In the present study the averages for maximal attention spans were not so high; for Combined Materials the average was 393.23 seconds; for all materials (except Companions) it was 327.62. It seems probable that this difference was due to the fact that many interest spans in this study were broken by distractions of five seconds which would have been ignored in the other investigation.

The longest time spent by any child with one toy, according to Herring and Koch, was 2202 seconds, the next longest interest span was 1655 seconds and the next 1230 seconds. Again the difference in recording time will explain the shorter length of interest span found in the observations of this investigation. Another explanation also is made in the first conclusion reached by Herring and Koch (13). "Interest span, as measured by our technique, was revealed to be a function of the type of toy offered for play, and the age of the child." The children demonstrating such long interest spans were four-year olds. The child who

stayed with Clay 1655 seconds was three years and one and a half months old. The other children were under four also, and their attention spans would, according to the conclusion quoted, be less than those of four-year olds.

It is interesting to consider the first part of the conclusion, and see how much the type of toy offered for play influenced the interest span in this investigation. Checking the data by dividing it shows consistency in play materials except Combined Materials. This varied from second to fourth place. But when it is remembered that this situation included every other material and thus gave stimulus for short or long spans, it is evident that variation in its rating in no way disqualifies the data or disagrees with the statement that interest span is found to be a function of the type of toy provided. It is interesting, in this connection, that the longest interest span recorded was in the combined situation and was with Clay.

When the interest span of the four distinct types of play material, Clay and Crayons, Blocks, Pictures and Books, and Housekeeping Materials (Table 10), is compared with their holding power (Table 2), there is agreement except in the case of Pictures and Books.

A picture book was used in the Herring-Koch study and the average interest span was also short. Their explanation that the book offered definite limits in use that the other play materials did not, therefore limiting the interest span, would also hold in the present investigation.

It appears that materials having the most "holding

power" throughout several observations usually stimulate the longest interest spans, but this may be altered, as with Pictures and Books or Combined Materials, because of the type and characteristics of the play material.

Extraneous Attention. Observers of the spontaneous play of preschool children, particularly Blatz and Bott, and Herring and Koch, have recognized the frequency of distraction in interest and attention. These moments of distraction have been arbitrarily defined by duration in time, a method which seemed practicable in their investigations. However, as these authors recognize, the mental flow is continuous though varying, just as the flow of activity is, and its variations must be marked off by arbirtrary means. In the play situations presented to the children of this study, distractions, in the usual sense of the term, were entirely eliminated. The privacy of the observation room was strictly observed, and the window was above the children's heads. The play objects in the experimental set-up were definitely limited in type and number and their arrangement was stereotyped. Yet this careful provision of objects to be attended to and prevention of outside disturbance did not restrict the children's range of attention as much as might be expected. By the end of the experiment absolutely everything in the room within reach had been touched and otherwise used. The radiator guard had been climbed, the walls patted, the floor molding rubbed, the pattern of the door traced, the lights turned off and on, and so on indefinitely.

The number of minutes spent in attending to and manipulating things other than those included in the set-up of the situation were differentiated in the records and time graphs and the percentage of this "extraneous attention" was computed for each material (Table 11).

TABLE 11

Percentage of Time Spent with Extraneous Things

Rank (according to total time with material)	Material	Percentage of total time with material	Percentage of total time with extraneous things
I	Combined materials	94	6
II	Clay and crayons	90	10
III	Blocks	89	11
IV	Pictures and books	84	16
V	Housekeeping materials	82	18
VI	Companions	60	40
Average percentage of all materials		83	17
Average percentage of all materials (except companions)		88	12

As was to be expected, the empty room gave an extremely high percentage of time spent with extraneous things (40 per cent), while the percentage was low when many playthings were provided (6 per cent). The difference between these two extremes is particularly evident. The ranking of materials according to the amount of time they held the children's attention remained consistent when the data was split.

A comparison of this ranking with the "holding power" as measured by the total number of minutes spent with each material (Table 2) shows interesting similarities. The comparative rank of the materials in Table 11 agrees with that in Table 2 except for Pictures and Books. The drop in rank for this situa-

tion is explained by the shorter interest span revealed by the figures in Table 10. The same characteristic of this type of material to set a limit to the interest span by a definite mark of completion also freed the attention more frequently for extraneous objects.

The presence and type of this activity with extraneous objects is clearly enough defined throughout all the observations to warrant the conclusion that the children were sensitive to most of the details in their immediate environment and that, in the absence of an adult, they reacted to, and experimented with, any of these details. It is also evident that the amount of such activity is reduced by the providing of interesting play material as in Combined Materials and greatly increased by the lack of play material in a bare room.

Attention Given to Companion. One quality which is desirable in play materials for young children, who are in the midst of acquiring social attitudes, is that of inducing cooperative play. Since there was very little friction between the children during the observation periods, the number of minutes of attention given to the

TABLE 12
PERCENTAGE OF TIME SPENT IN COMPANION PLAY

Rank	Material	Percentage of total time spent attending to companion
I	Companion	27
II	Combined materials	18
III	Blocks	17
IV	Housekeeping materials	15
V	Clay and crayons	14
VI	Pictures and books	12
	All materials	17
	All material (except companions)	15

companion provides a measure of the amount of social play. The percentages are given in Table 12.

The percentage of time spent in attending to or playing with the companion is noticeably small with every material and is in harmony with the emphasis often put upon the individualistic character of the play of two- and three-year olds. The situation called Companions, as was anticipated, forced the percentage to its highest, but 27 per cent is not enough really to warrant the title. The amount of Companion play with each material shows no relation to the children's preference for materials or to any of the other characteristics measured. However, the materials showed consistency in the maintenance of rank except in the case of Blocks which varied from 13 to 21 per cent when the data was split.

Attention to Self. It is revealing that in making the time graphs it was not only necessary to include everything in the room such as the door, lights, and radiator, but also necessary to make a column for "self." The infant's play with his own body and his partial knowledge of and mastery over himself through this play during the first year has been portrayed by Charlotte Bühler (4). The two- and three-year olds of this study were still interested in "self play," performing somewhat complicated feats such as galloping, rolling on the floor, and jumping. This was a small part, at least, of the play activity of each situation shown in Table 13.

Division of data on self play disclosed some consistency within the materials. The extremes were definite and widely differentiated, Companions drawing 32 per

TABLE 13

PERCENTAGE OF TIME SPENT IN ATTENTION TO SELF

Rank	Material	Percentage of time
I	Companions	31
II	Pictures and books	6
III	Blocks	4
IV	Housekeeping materials	4
V	Clay and crayons	3
VI	Combined materials	2
	Average percentage with all materials	8.6
	Average percentage with all materials (except companions)	3.8

cent and Combined Materials 2 per cent of self play, while no toys increased the percentage about seven times more than the average amount with playthings. Pictures and Books have the highest percentage of self play for any single play materials. These percentages for companion play and self play indicate the predominance of play with materials. The graphs have pictured the tendency for this play of the two children to be with similar materials at the same time, but with few overt signs of attention to one another. Table 14 presents a comparison of the percentages of these dif-

TABLE 14

DISTRIBUTION OF TIME IN SOCIAL-PERSONAL PLAY

Material	Percentage of time in self play	Percentage of time in companion play	Percentage of time in parallel play
Combined materials	2	18	70
Clay and crayons	3	14	83
Pictures and books	6	12	82
Blocks	4	17	79
Housekeeping materials	4	15	81
Companions	31	27	42
All materials	8.33	17.1	74.5
All materials (except companions)	3.8	15.2	81

ferent types of play: self play, companion play, and play with materials or parallel play.

A study of this comparison shows a general tendency for parallel play to predominate and for companion play to be greater than self play. Yet there is a distinct variation from this in the case of Companions.

D. Comparative Evaluations

Throughout this phase of the study of the various aspects of the children's reactions with play materials, differences have been noticeable and the preceding tables afford a basis for evaluation and show similarities as well as differences in the play situations. In evaluating these situations briefly it may be said that Combined Materials offer the best possibilities, holding the children for the longest time, involving the greatest variety of activity, and including more than the average amount of companion play and only an average amount of non-acceptable activity.

Clay and Crayons are next in value, holding the children for the longest time, involving the second amount of variety in activity, surpassing Combined Materials in length of interest span, but encouraging little companion play and involving a greater percentage of non-acceptable activity.

Pictures and Books and Blocks each maintain a middle rank but have different virtues and shortcomings. While Pictures and Books seemed to have a more consistent holding power for the children this interest was limited by the nature of the materials. Play with books was quiet and slow and little of the activity was "bad."

Blocks, however, excelled all materials in ability of activity and in encouraging cooperative play as well as adapting to individual play.

Housekeeping materials had less holding power than the other materials. Activity in this situation was swift and only a small percentage was unacceptable. In other respects this type of play material possessed fewer virtues than the others.

"Companions" had but little holding power as compared to the other play situations. The variety and amount of activity in this set-up was small, compared to the others and the rapidity, repetition, non-acceptability and self-centeredness of the activity were unusually high.

IV

COMPARATIVE STUDY OF THE INDIVIDUAL CHILDREN

A. Activities and Distribution of Time

The analysis thus far has shown definite and consistent differences between the various types of play materials when the combined reactions of the children are considered. This study was planned to compare and evaluate play materials and was in no way a test of the abilities or development of the children. But the records offer information concerning each child's reactions just as they present the information which has been analyzed for each material. Considering the data from the standpoint of the children throws more light upon the behavior of children left to their own devices and the value of the play materials.

A tabulation of the number of minutes each child spent with the various materials is given in Table 15. This shows the materials maintaining ranks similar to those in Table 2 and also indicates individual preferences for materials as measured by the amount of time spent with them. This information presented graphically in Figure 3 displays more clearly the similarity to Table 2. Combined Materials are definitely the highest, while Companions are as decidedly the lowest. Clay and Crayons take an unquestionable second rank. Pictures and Books and the setting, Blocks, show the same similarity indicated by the number of minutes spent with them, and again Pictures and Books are somewhat lower. Also, in this graph Housekeep-

TABLE 15

Number of Minutes Each Child Spent with the Materials

Child Rank	Andrew	Barbara	Connie	David	Edward	Frank
I	Clay and Crayons 79	Combined Materials 143	Combined Materials 188	Combined Materials 173	Clay and Crayons 97½	Clay and Crayons 155½
II	Pictures and books 70	Clay and Crayons 101	Clay and Crayons 117	Clay and Crayons 123½	Combined Materials 92½	Combined Materials 145
III	Blocks 66	Blocks 85	Pictures and books 102	Blocks 81½	Housekeeping Materials 80½	Pictures and books 102
IV	Housekeeping Materials 57½	Pictures and books 68½	Blocks 88	Pictures and books 64	Pictures and books 72½	Blocks 76
V	Combined Materials 45	Housekeeping Materials 67	Housekeeping Materials 65	Companions 63½	Blocks 65	Housekeeping Materials 64
VI	Companions 34	Companions 42	Companions 62½	Housekeeping Materials 47	Companions 28	Companions 57½

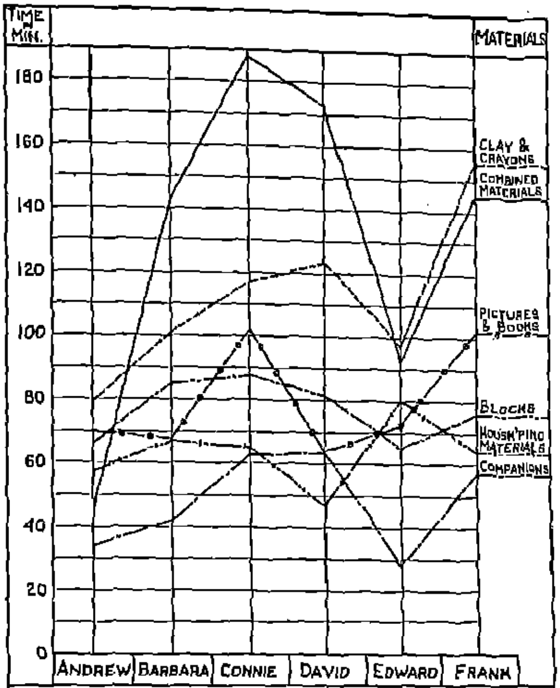


FIG.-3
TOTAL NUMBER OF MINUTES EACH TYPE OF MATERIAL WAS USED BY EACH CHILD

ing Materials rank low. These differences are definite, but the lines of the graph are not straight and the peaks indicate variations in the children's reactions.

The data from Table 15 presented differently in Table 16 show the individual children ranked according to the time they spent with each material. In Figure 4 these data are presented graphically. Much more similarity is evidenced in the trend of these lines than in those lines representing materials plotted in Figure 3. There is more frequent overlapping of the lines

TABLE 16

NUMBER OF MINUTES EACH CHILD SPENT WITH THE MATERIALS

Material	Combined materials	Clay and crayons	Pictures and books	Blocks	Houskeeping materials	Companions
Rank I	Connie 188	Frank 155½	Connie 102	Connie 88	Edward 80½	David 63½
II	David 173	David 123½	Frank 102	Barbara 85	Barbara 67	Connie 62½
III	Frank 145	Connie 117	Edward 72½	David 81½	Connie 65	Frank 57½
IV	Barbara 143	Barbara 101	Andrew 70	Frank 76	Frank 64	Barbara 42
V	Edward 92½	Edward 97½	Barbara 68	Andrew 66	Andrew 57½	Andrew 34
VI	Andrew 45	Andrew 79	David 64	Edward 65	David 47	Edward 28

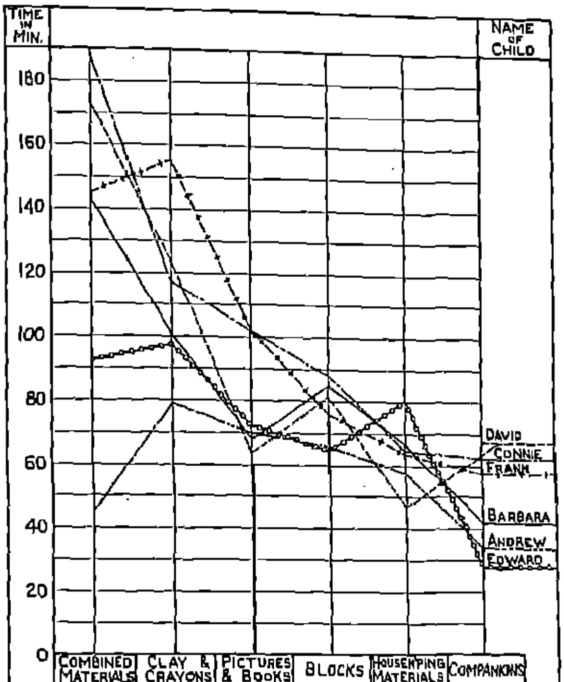


FIG.-4
TOTAL NUMBER OF MINUTES SPENT WITH EACH TYPE OF MATERIAL AS OBSERVED

representing time spent *by* the children than there is in those spent *with* the materials. These two graphs show a consistent trend in the time reaction of the children as influenced by the controlled variations in unchanging materials. There are wide differences between the selected materials and narrow differences, resulting in much similarity, between the selected children. However, individual differences are apparent, for instance in the lines plotted for Connie and Andrew.

An analyzation of the records from the standpoint of each child was made with tables for (1) total number of recorded activities, (2) rapidity of activities, (3) percentage of repetition in activity, (4) variety of activity, and (5) socially acceptable activity. Also the (1) total time spent with materials (graphed in Figure 6), (2) average and range of length of observations, (3) average length of interest span for each child, (4) percentage of time spent with extraneous things, (5) percentage of time each child spent in companion play, (6) in attention to self, and in (7) parallel play. The ranking of the children in each of these Tables[8] was checked by halving the data and in no case was consistency in ranking maintained when the substitute for a repetition of the experiment was made. In the same consideration of the playthings there was consistency of rank maintained in (1) the amount of time spent with them, (2) in the social acceptivity of the activity, (3) in the amount of time given to extraneous things, (4) in the time given to companions. There was identity of ranking in the materials, also, when variety of activity was compared to the total amount of time spent with materials. It is evident that the ranking is not static for ever-changing children as it is for unchanging materials.

The data revealed much similarity in the children's reactions; for instance, the range in number of activities per minute was from 4.65 to 4.96 as compared to the range between 4.00 and 5.72 for materials. Rank

[8]Tables available in complete study filed in the Library of Yale University.

according to percentage of repetition in activity placed the children's names in chronological order. This was interesting, but the data were not sufficient to justify conclusions on the decrease of repetitive activity as age increased. The amount of repetition in the children's activity ranged from 55 per cent to 64 per cent as compared to 55 per cent to 66 per cent in the data for materials, showing again that more than half of the unguided play activities of children are repetitive no matter who the child or what the situation. More similarity in length of interest span was shown. The average for children ranged from 70.5 seconds to 88.5 while the same data from the standpoint of materials showed the range in average length of interest span to be from 45.53 seconds to 115.63 seconds. The information for percentage of time spent in attention to extraneous things also shows a smaller range for children, 10 per cent to 20 per cent as compared to the difference between 6 per cent and 10 per cent given in Table 11. To complete the analysis the percentage of time each child spent in attention to self, in companion play, and in parallel play was tabulated and portrayed the same general pattern as that for materials given in Table 14.

B. The Influence of Companions

A further consideration in the analysis of the social play of these children is the effect of one individual upon another. Did a child remain in the observation room longer when paired with a certain other child, no matter which play situation was offered? The

children's preferences for each other can also be measured by the time spent in the room together. This measure was made and indicated by a table for each of the six children in which the other five children were ranked according to his preference as measured by time. Again the data was divided in two different ways to check the consistency of preference thus measured. No child indicated a definite order of preference for the other children. A graph of the total number of minutes each child spent with his companions

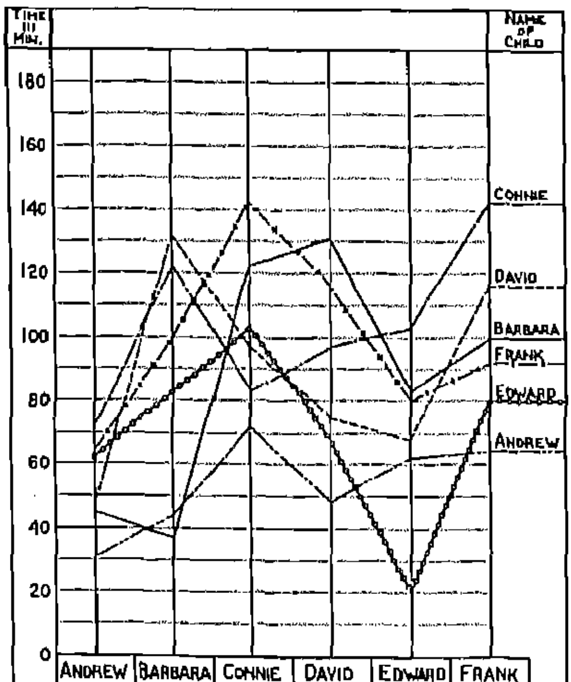


FIG.-5
TIME SPENT BY EACH CHILD WITH EACH COMPANION

(Figure 5) shows neither the wide differentiation between lines indicated by the same method of graphing the total number of minutes spent with each material (Figure 3) nor the similarity in the trend of lines indicated in the graph of the number of minutes each individual spent in the observation room (Figure 4). This graph and the data which it illustrates lead to the conclusion that there was little influence of one individual upon another. This was to be expected since the children were similar in many respects and were thoroughly accustomed to playing with one another in the Guidance Nursery. And these facts lead to the further conclusion that, since the individuals were similar and were paired in the same variety of ways throughout each different play situation presented, the plan of procedure made each play situation comparable except in the change of type of material offered.

C. Individual Differences

The foregoing analysis of the activity and distribution of time for the individual children, when compared to the similar analysis for materials, emphasizes the fact that the materials were intentionally variables, unknown variables, and that, being subjected to the same six children under the same conditions they showed consistency in their variations; and, being inanimate, their functions in the play situations were unchanging from observation to observation.

The analysis shows the children following a general play pattern with similar reactions changing in similar

ways with the change of play materials. But the reactions, though patterned, were not stereotyped, and unlike the inanimate play things, the children's functions in the play situations were different and adaptive in varying degrees from one observation to another. The lines representing materials, in Figure 3, though clearly differentiated are not smooth. Figure 4 shows a similar trend in the lines representing the children but individual differences are apparent. For instance, Connie's line is outstanding, showing that she tended to remain in the observation room for comparatively long periods of time. A study of the aforementioned tables shows that she had the greatest total number of minutes spent with materials, also the longest average length of observations, and the longest average interest span of 88.5 seconds. In each of these cases she maintained first or second rank when the data were divided. She also evidenced a greater variety in activity than the other children. The ability to remain happily occupied without supervision and the tendency to attend to things an unusually long time is in agreement with her mother's report of her home play and training leading to self-reliance. The records of the Clinic of Child Development also indicate this tendency consistently. The Guidance Worker recorded her first impressions when Connie entered the Guidance Nursery at 21 months of age as, "She was unusually independent of adults, was habitually contented and happy."[4] This same tendency is noted more definitely in the examination made when Connie was three years

[4]From Miss Ann Jennings' records.

old—"Child never seems to have enough and would have liked to continue longer even at the end of three-quarters of an hour. Very good attitude and control."[5]

In these characteristics Connie differed in degree from the other children but she remained in the middle ranks so far as rapidity of activity, repetition of activity, and the amount of time with extraneous things and with companions were concerned. In all these considerations, it will be remembered, the children are very similar. However, she has the second highest percentage of non-acceptable activity, 9.5 per cent. She is second highest in the percentage of self-play which accounts for her having a comparatively smaller percentage of parallel play.

The analysis of the children's reactions has shown them to be quite similar and their comparative ranks have shifted frequently when halves of the data were compared. A study of the findings concerning the four children, Barbara, David, Edward, and Frank, who maintained the more intermediate ranks discloses more similarity than difference and numerous shifts in position. Even so the records do not conceal individual variations. For instance, divided data showed David varying between rank II and V or VI, yet the total figures for his reactions are, in most cases, above the average of the group. Throughout the observations it was clear to the examiner that David's tendencies to unusual variety and persistence in activity were often cut short because of his extreme sensitiveness and

---

[5]*From the record of Development* Examination made by Dr. Ruth W. Washburn.

reactiveness to sounds, the behavior of his companions, and other environmental factors. This caused unusual variation in his reactions and ranking. The existence of this sensitiveness in other situations is brought out over and over in the Guidance Nursery records, the Developmental Examinations, and parental interviews.

Frank, the youngest child in the group, is consistent in having the fewest recorded activities per minute and the smallest percentage of non-acceptable activities. The amount, variety, and length of time spent in play activities compare well with the older children and this is in harmony with other records of his slow-moving, quiet behavior yet superior ability. In the same way Barbara and Edward displayed characteristic differences, many examples of which could be taken from the observation records with similar characteristic reactions quoted from the records of the clinic and nursery.

However, the child showing the most noticeable difference in response to the play materials was Andrew, the oldest in the group. His line in Figure 5 is the one which differs most obviously from the trend of the others. As the clearest example of individual differences a more complete consideration of his reactions and the clinic records will be included. The information in the various tables of the preceding section shows that he spent the least amount of time in the observation room. His preference for play material agreed with the preference of the group except in that of Combined Materials which the group preferred most high-

ly, but which Andrew seemed to prefer little more than a bare room. As well as spending the fewest number of minutes in play, he has the smallest number of activities recorded. His activity was more rapid than that of most of the children, it included a consistently small amount of repetition, a consistently small amount of socially non-acceptable activity, and also the least variety.

In the division of time the average length of his play periods were, of course, the shortest. He spent the smallest percentage of time attending to extraneous objects, the next to the smallest percentage of attention to his companion or in self play. Because of the small amount of time spent in overt attention to the companion or in self play, he has the next to the greatest percentage of time in parallel play. He seemed to prefer the younger children, Edward and Frank, or sometimes Connie, as companions, though his preference is not marked. As a corollary to the small amount of time spent in the observation room there is the small amount of time spent with any one child.

The characteristics thus portrayed show interrelations with one another. Since Andrew's activity was more rapid and less repetitional than that of most of the children, and since he spent little time in attending to extraneous things or in simply running around in mere self play, it is natural that he should, as he often announced, get "all through" with dispatch. But interrelationships are not explanations. The records of the clinic offer these. From the time of Andrew's entrance into the group of the Guidance Nursery his in-

terest in people was noticeable. At the age of two and a half he was aware of the entry of a new child into the group and would always approach the child and usually attempt to hug him. He did this so consistently that the procedure was almost a recognized test situation for new children.

The records of his activity in the Guidance Nursery show growing interest in blocks and paints and clay, but they are more often concerned with Andrew's social play, citing instances of playing with or teasing younger children and attempting to play with older children and sometimes being repulsed. When he was two years and seven months old, the Guidance worker recorded that he "dominates younger children, but plays with older ones very well."[6] His interest in the group of children and the teacher appears throughout his records. When he was slightly older his attitude toward being alone is revealed in the following record (made January 16, 1930): "Was cold outdoors. Wanted to come in. Did not like to play alone indoors. Usually likes pencils and paper, but was not interested when left alone with them."[6]

The records made during the time of the experiment continue to stress his social relationships in play, telling of games played with older children who were temporary members of the group such as one in the jungle gym (February 17, 1931) when he said, "I was the moon and they were the two stars" (24). They tell of his directing the younger children and even later of

[6]From Miss Ann Jennings' records.

teasing a child more than a year his senior by statements of superiority.

A comparison of these records and the records of the experiment lead to the interpretative explanation that Andrew brought into the play situations an unusually social personality. He was always thoroughly cooperative with the examiner and willing to go and play. He followed the directions of "having a good time" and calling when he was through. Yet it was plain that he finished quickly and was glad to return to a group of children and to adults. Six times, also, his leadership, which was usually tactful, became so domineering that his companion called to leave. In this case, more outstanding perhaps than the others, individual differences in personality are seen to influence play reactions, interest in people shortening the play with things.

V

THE TOTAL PLAY PATTERN

There were individual differences and differences in the use of material yet similarities of reaction appeared in every situation in this experimental procedure like the repetition of various motifs in a complicated tapestry. These 226 observations with their records, graphs, and charts disclosed a pattern of tendencies in spontaneous play which, because of their consistent appearance over and over again, and their agreement with previous observations on play, are probably typical of the unguided play reactions of children of the mental level, social level, and limited age range which these represent.

The first question which was answered, and the one upon which the whole experiment rested, was whether or not the children would remain alone, or with each other, in the observation room long enough to afford a usable record of their behavior. In order to test the holding power of the different materials the observations were stopped at the first sign of boredom, or desire to be somewhere else, as indicated by a second call, or at the first hint of an emotional upset. The length of the observations measured the amount of time the children played contentedly (usually gayly). The total picture of spontaneous play showed the children's willingness to play without an adult's presence and their ability to do so for more than half an hour. (The preliminary study included one record of 55 minutes and another of an hour of happy, unguided play. There

was no nursery group of children to return to in these cases). However, the average time of play alone was 14 minutes and 12 seconds and the range in length of the observations was from one-half a minute to 41 minutes.

The reasons for stopping the periods of play sometimes quickly, sometimes after a comparatively long time, are familiar ones to parents and teachers of young children. The usual reason was that interest in the play material lagged and the child preferred to return to the teacher and other children. The majority of the observations, 87 per cent, were calmly stopped by the children announcing that they were ready to go. But in ten cases lonesomeness brought tears, and in three cases quarreling stopped the play. Even though every possible precaution was taken to prevent physical danger, play was stopped twice because the children got themselves into dangerous situations. Also all possible forethought was taken in the experimental procedure to avoid interference with, or by, the daily routine, yet 12 times observations were stopped because of routine matters. These latter observations, however, were unusually long ones.

In spite of these inevitable breaks, the one way vision screen disclosed a drama more original, swifter, and more amusing than any ever caught upon the silver screen. And the first outstanding fact impressed upon the recorder in the attempt to preserve this drama on paper was the continuousness and swiftness of activity. The figures will make this evident. In the 3069 minutes of observation 14,868 activities were recorded. This

gives an average of 4.84 activities per minute—a shift in activity every 12 seconds. It is further evidence that children can be depended upon to react in some way and that in play their reactions are continuous and swift.

So much moving about involves repetition of activity. Preschool children have often been characterized as "liking to do the same thing over and over." No doubt repetition is enjoyable to them and necessary to their learning. It was evident in every observation. A study of the activity charts showed that 59 per cent of the total number of spontaneous play activities were repeated activities. And that no matter who the child or what the material, more than half the activity was repetitional. This count did not take into consideration the way in which children of similar ages repeat one another's behavior. When this was accounted for also, it was found that 81 per cent of the activities recorded during the whole experiment were repetitive.

But the impression the experimenter received while recording was that of varied activity. Only 19 per cent of the activities were original, that is, appeared for the first time in this group. But this gave the actual count of 2821 different types of activities. A different kind of activity almost every minute is enough to give the impression of variety!

The social acceptability of this activity is probably of more practical interest than its rapidity or variety. The criteria for judging the children's behavior have been discussed (pages 43 to 47). Of the entire recorded activity 8 per cent was found to be non-accept-

able. When only the situations which provided play materials are considered, 5.2 per cent of the activity was non-acceptable and even with the most stringent checking of annoying as well as "bad" behavior throughout every situation it was found that only 11 per cent could be checked, leaving 89 per cent of the children's behavior which was unquestionably acceptable and "good."

These characteristics of play behavior were revealed in the activity charts. The time graphs presented characteristics of the division of time in play. The examples given in Figures 1 and 2 show the way in which shifted attention or interest followed a pattern of shorter and longer intervals. These graphs picture the similarity in the children's distribution of time and also show individual differences. They show clearly the difference in pattern between Housekeeping Materials, having short interest spans, and Clay and Crayons having long interest spans.

When activity is so rapid and varied the length of interest spans must be short. The average length, in situations which provided play materials, was 81 seconds and for all of the situations, including Companions, was 76 seconds. While children's interest shifts often, it is not always brief. Many times the span of interest lasted only five seconds, but many times it was longer than the average. The range was great. The longest interest span recorded was 1655 seconds.

It will be remembered that one reason for the entry of small units of time in the records was the tendency of the children to attend briefly to other things about

the room rather than the play materials provided. Because of this it was necessary to include everything in the room in the time graphs; the door, the radiator guard, and so on. It was found that an average of 17 per cent of the time was spent in attention to extraneous things. It was a very definite part of the play in each situation though it was decreased or increased by a change in play material. A realization of the existence of this all-inclusive interest and activity helps in understanding how children learn in all environments, fall into danger in thoughtlessly arranged environments, and are accused of "meddling" and made unhappy in environments which are inconsiderately planned. That the play materials held the center of interest, according to the plan of procedure, is shown by the fact that they were given 83 per cent of the time.

These young children were not only interested in things, they still showed an interest in "self play" which in infancy had been all absorbing to them. This activity involved neither the materials nor the companion but was merely a capering or tumbling or running about which appeared in the time graphs as solid blocks rather than being scattered throughout the observation. It occurred frequently, but made up only 8.6 per cent of the total activities recorded. And when only situations which provided play materials were considered it made up 3.8 per cent of the total time.

Children of these ages are also interested in other people. They enjoy and need the companionship of

their peers. In the columns of the graphs labeled "Companion" were recorded the actual moments of attention, as judged by overt behavior, given to the other child in the room. This was nearly twice as much as that given to self, but was only 17 per cent of the total time.

The outstanding characteristic of this "companion play," as shown in the time graphs, contrasted clearly with "self play." Attention to the companion usually lasted for only a few seconds at a time and these moments were scattered throughout the whole period of observation, while self play usually appeared in one single block of time. To the experimenter it appeared that the companion was usually in the periphery of attention and became the center of attention only during the few seconds indicated by bits of overt behavior which were the beginning of real cooperative play and consideration of others. The self play seemed to be "time out," a return to an earlier, easier type of play and a relaxation from the more taxing creation and experimentation.

But the most outstanding characteristic of the division of time in this spontaneous play was that the children were usually doing similar things with similar materials at the same time, but gave no evidence of attention to each other. This has been designated "parallel play" and recognized as a characteristic of the spontaneous play of children of preschool age (12, pp. 258-259). The children under observation devoted 74 per cent of their time to this type of play.

VI

CONCLUSIONS

1. The first definite accomplishment was a method of recording the observed activity and tabulating it in an easily comparable way. The tabulation was in the form of time graphs (pp. 22, 23, 27) and activity charts (pp. 29-32).

2. The materials, when tested and compared by the experimental procedure, manifested differences. Considering the data as a whole it could be said that Combined Materials offered the best play situation; that Clay and Crayons ranked highest, in general, of the single types of play materials; that Pictures and Books and Blocks each held an intermediate rank; while Housekeeping Materials were lower, and Companions were decidedly the poorest play situation for the children observed. The play materials possessed different qualities and a summary is given on page 58.

3. The holding power of a material was measured by the length of time the children voluntarily remained with it in the examination room. One factor which determined the length of time a material would be used was the variety of activities to which it lent itself in the children's play. The materials ranked according to the children's preference or "holding power" and variety of activities were:

I Combined Materials
II Clay and Crayons
III Pictures and Books
IV Blocks

V Housekeeping Materials
VI Companions

4. It was evident that a change in play environment influenced the play reactions of the children studied. This was most clearly brought out in the contrast between the situations called Combined Materials and Companions. The children remained 786½ minutes with Companions. There were 829 different activities with Combined Materials and only 276 different activities with Companions. Only 7 per cent of the activity with many playthings was socially non-acceptable while 21 per cent of the activity in a bare room was non-acceptable. The average length of interest span for combined materials was 94 seconds but in the bare room the average interest span was only 45 seconds long. In a bare room the children spent 40 per cent of their time with "extraneous things." When many play materials were provided only 6 per cent of the time was given to extraneous things. With Companions 27 per cent of the time was spent in attending to the playmate while 18 per cent of the time in the Combined Materials was spent in attending to the playmate. But with Companions 31 per cent of the time was spent in self-play while the room providing many toys had only 2 per cent of the time in attention to self.

5. These last percentages indicate that depriving these children of play materials forced them much more to attend to themselves than to play with their companion. This is in agreement with the term "indi-

vidualistic" often applied to the play of children within this age range.

6. The children, when compared in the same way, showed no such consistent differences as the materials. However, consistencies within each child's reactions indicated that individual differences, particularly personality differences, were an influencing factor in the play reactions observed.

7. A study of the children's preferences for each other showed no consistent influence of one individual upon another. This added to the fact that the children were similar and were paired in the same variety of ways throughout each play situation gives further assurance that the plan of procedure made each play situation comparable except in the change of type of play material offered.

8. The entire data revealed a total pattern of play with play reactions showing similarities in each situation yet changing in degree as materials were changed or following individual differences.

The outstanding characteristics of these play reactions were: swiftness, an average of 4.84 activities per minute; repetition, an average of 59 per cent; and variety, with 2821 different types of activity recorded in 3069 minutes. Of all this activity only 8 per cent was socially non-acceptable, but this amount changed consistently with the change of play material.

The time of play was distributed in a patterned way. The average interest span was found to be 81 seconds. Eighty-three per cent of the time was given to the material provided, but 17 per cent of the time was given

to extraneous things and the children were inclusive in their interests showing that they were aware of everything within reach and sight in the play room. Self-play took only 8.6 per cent of the total time, but was a definite part of the children's reactions. The companion was given 17 per cent of the time, but the greatest amount, 74 per cent was devoted to parallel play.

9. The children were willing to play alone and able to do so for more than half an hour. However the average time which elapsed before they needed or wanted adult guidance was 14 minutes and 12 seconds. The range in length of observations was from one-half a minute to 41 minutes.

References

1. Blatz, W. E., & Bott, H. Parents and the preschool child. New York: Morrow, 1930. Pp. 352.
2. Bott, H. Observation of play activities in a nursery school. *Genet. Psychol. Monog.*, 1928, 4, 44-88.
3. Bridges, K. M. B. Occupational interests of three-year-old children. *Ped. Sem.*, 1927, 34, 415-423.
4. Bühler, C. The first year of life. New York: Day, 1930. Pp. 281.
5. Cowan, A., & Carlson, A. D. Bringing up your child. New York: Duffield, 1930. Pp. 278.
6. Cushing, H. M. A preservation tendency in preschool children, a study of personality differences. *Arch. Psychol.*, 1919, 17, No. 108. Pp. 55.
7. Foster, J. C., & Mattson, M. L. Nursery school procedure. New York: Appleton, 1929. Pp. xiv+220.
8. Freeman, F. N., Holzinger, K. V., & Mitchell, B. C. The influence of environment on the intelligence, school achievement, and conduct of foster children. *27th Yrbk. Nat. Soc. Stud. Educ.*, Pt. 1, 1928, 102-217.
9. Gesell, A. L. Infancy and human growth. New York: Macmillan, 1928. Pp. 418.
10. ———. Guidance of mental growth in infant and child. New York: Macmillan, 1930. Pp. xi+322.
11. Gesell, A. L., & Thompson, H. Learning and growth in identical infant twins: an experimental study by the method of co-twin control. *Genet. Psychol. Monog.*, 1929, 6, 1-124.
12. Goodenough, F. L., & Anderson, J. E. Experimental child study. New York: Century, 1931. Pp. xv+544.
13. Herring, A., & Koch, H. L. A study of some factors influencing the interest span of preschool children. *J. Genet. Psychol.*, 1930, 38, 249-279.
14. Hulson, E. L. An analysis of the free play of ten four-year-old children through consecutive observations. *J. Juv. Res.*, 1930, 14, 188-208.
15. Johnson, H. M. Children in the nursery school. New York: Day, 1928. Pp. xx+324.

UNE ÉTUDE DU JEU DES ENFANTS D'ÂGE PRÉSCOLAIRE PAR UN OBSERVATEUR NON OBSERVÉ

(Résumé)

Pour contribuer plus à la compréhension du jeu et des besoins du jeu des jeunes enfants, on a fait cette expérience dans le but de montrer: (1) si un changement du milieu du jeu influait sur les réactions des enfants étudiés; (2) s'il en était ainsi, quelles différences se montraient dans les réactions aux divers matériaux présentés; (3) jusqu'à quel point ces enfants devraient ou pouvraient être permis à suivre leurs propres inventions; et (4) de trouver d'autres renseignements sur les caractéristiques générales ou individuelles des réactions dans le jeu montrées par les enfants observés.

L'expérience consistait à voir chacun de six enfants d'âge préscolaire mis en paire avec tout autre enfant dans le groupe, et seul, dans six milieux différents du jeu. Ces milieux ont été faits dans une salle d'examen contenant un écran permettant la vision dans une direction seulement, lequel a permis l'enlèvement de l'influence immédiate adulte mais a rendu possible une étroite observation. Les situations de jeu présentées ont été: des Matériaux de ménage, des Blocs, des Tableaux et des Livres, de l'Argile et des Crayons de pastel, des Compagnons (nuls matériaux du jeu dans la salle), et des Matériaux combinés (quelques-uns de chacun des quatre premiers types de matériel inclus.)

Les enfants ont été entièrement libres de jouer comme ils voulaient et de partir quand ils le voulaient. On a enregistré leur comportement immédiatement dans un compte courant. Les points notés ont été: (1) le matériel approché; (2) le temps de l'approche (au plus proche cinq secondes); (3) les changements de l'activité pendant qu'ils étaient avec le matériel; et (4) le temps du changement de l'activité. Ces notations ont été organisées et disposées en tables en forme de graphiques de temps et tables d'activité, de sorte que l'activité et la distribution du temps dans toute observation ont été clairement et quantitativement comparables à celles de n'importe quelle autre.

Les principaux résultats de cette expérience ont été les suivants:

1. Il a été évident qu'un changement du milieu du jeu a influé sur les réactions du jeu des enfants étudiés. Ceci s'est montré le plus clairement dans le contraste entre les situations appelées "Matériaux combinés" et "Compagnons." Les enfants sont restés 786½ minutes avec les "Matériaux combinés" comparé à 287½ minutes avec les "Compagnons." Il y avait 829 activités différentes avec les "Matériaux combinés" et seulement 276 activités différentes avec les "Compagnons." Seulement 7% de l'activité avec beaucoup de jouets a été socialement non acceptable, tandis que 21% de l'activité dans une salle dégarnie a été non acceptable. D'autres contrastes se sont montrés dans l'étude.

2. L'enlèvement des matériaux de jeu de ces enfants les a forcés beaucoup plus à s'occuper d'eux-mêmes qu'à jouer avec leurs compagnons.

3. Les différences individuelles, surtout les différences de personnalité, ont été un facteur d'influence dans les réactions du jeu observées.

4. Toutes les données ont montré une forme totale du jeu avec des réactions du jeu qui montraient des similarités dans chaque situation mais changeaient de degré avec le changement des matériaux ou suivaient les différences individuelles. Quelques-unes de ces caractéristiques ont été: la rapidité, une moyenne de 4.84 activités par minute; la répétition, une moyenne de 59%; et la variété, avec 2821 différents types d'activité en-

registrés en 1069 minutes. De toute cette activité seulement 8% a été non acceptable. Le temps moyen s'écoulant avant que les enfants aient eu besoin de la direction des adultes ou l'aient voulue a été de 14 minutes et 12 secondes.

COCKRELL

EINE UNTERSUCHUNG DER SPIELTÄTIGKEIT BEI VORSCHULPFLICHTIGEN KINDERN VON EINEM UNGESEHENEN BEOBACHTER

(Referat)

Um zu dem Verständnis des Spieles und der Spielbedürfnisse junger Kinder etwas beizutragen, wurde dieses Experiment ausgeführt, um die folgenden Punkte nachzuweisen: (1) ob eine Veränderung der Spielumgebung die Reaktionen der untersuchten Kinder beeinflusste; (2) wenn ja, welche Unterschiede in den Reaktionen auf die verschiedenen dargebotenen Materialien vorkamen; (3) den Grad, zu dem die Kinder allein bleiben sollten oder konnten; und (4) um noch mehr Auskunft über allgemeine oder individuelle Eigenschaften der Spielreaktionen zu finden, wie sie bei den Kindern beobachtet wurden.

Jedes der sechs Kinder des vorschulpflichtigen Alters wurde mit jedem anderen Kind der Gruppe gepaart, allein gelassen, und in sechs verschiedene Spielsituationen gebracht. Diese spielsituationen fanden in einem Untersuchungszimmer statt, das einen Gesichtschirm besass, der die ungesehene Beobachtung des Kindes gestattete. Die Spielsituationen waren Haushaltungsmaterialien, Blöcke, Bilder und Bücher, Ton und Kreidestifte, Spielgenossen (kein Spielmaterial im Zimmer) und verbundene Materialien (einiges von jedem der vier Typen von Material).

Es stand den Kindern frei, zu spielen, wie sie wollten und zu gehen, wenn sie wollten. Ihr Verhalten wurde dauernd aufgeschrieben. Die Einzelheiten, die bemerkt wurden, waren (1) das gebrauchte Material; (2) die Zeit der Annährung (bis zu fünf Sekunden); (3) Verschiebung der Handlung, während sie sich mit dem Material beschäftigten; und (4) die Zeit der Verschiebung der Handlung. Diese Ergebnisse wurden zusammengebracht und in der Form von Zeitzeichnungen und Handlungstabellen aufgestellt, so dass die Handlung und Verteilung der Zeit in irgendeiner Beobachtung klar und quantitativ miteinander vergleichbar waren.

Die Hauptbefunde dieses Experiments waren:

1. Es stellte sich heraus, dass eine Veränderung in der Spielumgebung die Spielreaktionen der untersuchten Kinder beeinflusste. Dies zeigte sich am klarsten im Gegensatz zwischen den Situationen, die "verbundene Materialien" und "Spielgenossen" genannt wurden. Die Kinder blieben 786½ Minuten mit den "verbundenen Materialien" im Vergleich mit 287½ Minuten mit den "Spielgenossen." Es gab 829 verschiedene Handlungen mit "verbundenen Materialien" und nur 276 verschiedene Handlungen mit "Spielgenossen." Nur 7% der Tätigkeit mit vielen Spielzeugen war sozial unannehmbar, während 21% der Tätigkeit in einem blossen Zimmer unannehmbar war. Weitere Gegensätze zeigten sich in dieser Untersuchung.

2. Wenn die Kinder ohne Spielmaterialien waren, beschäftigen sie sich mit sich mehr als mit ihren Spielgenossen.

3. Individuelle Verschiedenheiten, besonders Persönlichkeitsverschieden-

heiten, waren ein beeinflussender Faktor bei den beobachteten Spielreaktionen.

4. Alle Ergebnisse zeigten ein ganzes Gebilde des Spieles mit Spielreaktionen, die Ähnlichkeiten in jeder Situation aufwiesen, jedoch änderten sie sich im Grade, wie die Materialien verändert wurden. Einige dieser Eigenschaften waren: Schnelligkeit, im Durchschnitt 4,84 Handlungen pro Minute; Wiederholung, im Durchschnitt 59%; und Verschiedenheit, mit 2 821 verschiedenen Arten von Handlung in 3 069 Minuten. Von dieser Tätigkeit war nur 8% unannehmbar. Die durchschnittliche Zeit, die verging, bevor die Kinder die Lenkung der Erwachsenen brauchten oder wollten, war 14 Minuten und 12 Sekunden.

COCKRELL